A NOTE ON MEASURE FUNCTIONS IN A LATTICE[1]

M. F. SMILEY

We give first an equivalent statement of the measurability criterion of Carathéodory[2] which is applicable to an arbitrary lattice. We then study the closure with respect to finite and denumerable sums and products of the subset of measurable elements of a *modular* lattice. The case of regular[3] "outer measure functions" is then briefly discussed. The elements of the theory of lattices are presupposed.[4]

Let us consider a lattice L on which is defined a real-valued function $\mu(a)$. The elements $a \in L$ which satisfy

$$\mu(a+b)+\mu(ab)=\mu(a)+\mu(b) \tag{1}$$

for every $b \in L$ will be called μ-measurable. The totality of μ-measurable elements will be denoted by $L(\mu)$.

REMARK 1. *If L is a Boolean algebra and $\mu(0)=0$, then $a \in L(\mu)$ if and only if $a \in L$ and satisfies the condition of Carathéodory,[5] that is,*

$$\mu(b)=\mu(ab)+\mu(b-ab) \tag{2}$$

for every $b \in L$. For, if $a \in L$ satisfies (1), *the equation* (1) *and*

$$\mu(a+(b-ab))+\mu(0)=\mu(a)+\mu(b-ab)$$

yield (2). *The converse is proved by Carathéodory.*[6]

THEOREM 1. *If L is a modular lattice, then $L(\mu)$ is a sublattice of L.*

PROOF. Let $a, c \in L(\mu)$, $b \in L$. We obtain successively

$$\begin{aligned}\mu(a+(c+b))+\mu(a(c+b)) &= \mu(a)+\mu(c+b)\\ &= \mu(a)+\mu(c)+\mu(b)-\mu(cb)\\ &= \mu(a+c)+\mu(b)+\mu(ac)-\mu(cb).\end{aligned}$$

Since $c \in L(\mu)$ we have

[1] Presented to the Society, September 5, 1939.

[2] *Vorlesungen über Reelle Funktionen*, 2d edition, p. 246.

[3] Ibid., p. 258.

[4] See, for example, G. Birkhoff, *On the combination of subalgebras*, Proceedings of the Cambridge Philosophical Society, vol. 29 (1933), pp. 441–464; O. Ore, *On the foundations of abstract algebra* I, Annals of Mathematics, (2), vol. 36 (1935), pp. 406–437. The terminology and notation are those used by L. R. Wilcox and the author, *Metric lattices*, Annals of Mathematics, (2), vol. 40 (1939), pp. 309–327.

[5] Op. cit., p. 246.

[6] Ibid., p. 252.

$$\mu(c + a(c + b)) + \mu(ac) = \mu(c) + \mu(a(c + b)),$$
$$\mu(c + (a + c)b) + \mu(cb) = \mu(c) + \mu((a + c)b).$$

Using the modular law we see that $(a+c)(c+b)=c+a(c+b)=c+(a+c)b$. It is then clear that

$$\mu(ac) - \mu(cb) = \mu(a(c + b)) - \mu((a + c)b),$$

and (1) with a replaced by $a+c$ follows easily. Thus $a+c \in L(\mu)$. By duality, $ac \in L(\mu)$. This completes the proof.

DEFINITION 1. *If, for each increasing (decreasing) sequence* $(a_i; i=1, 2, \cdots)$ *of elements of* $L(\mu)$ *with a sum (product)* $a \in L$, *we have* $\lim \mu(a_i)=\mu(a)$ *as* $i\to\infty$ *we say that*[7] L *satisfies* $B^+(\mu)$ $(B^-(\mu))$; *if moreover* $\lim \mu(a_i+b)=\mu(a+b)$ *and* $\lim \mu(a_ib)=\mu(ab)$ *as* $i\to\infty$ *for each* $b \in L$, *we say that* L *satisfies* B^+ (B^-).

REMARK 2. *If* L *satisfies* B^+ (B^-), *then* L *satisfies* $B^+(\mu)$ $(B^-(\mu))$. *It suffices to take* $b=a$ *in the definition of* B^+ (B^-).

We shall assume throughout the remainder of this note that L is modular and that $\mu(a)$ is monotone increasing.

THEOREM 2. *A sufficient condition for closure of* $L(\mu)$ *with respect to denumerable sums (products) in* L *is that* L *satisfy* B^+ (B^-). *This condition is necessary if* L *satisfies* $B^+(\mu)$ $(B^-(\mu))$.

PROOF. To show that B^+ is sufficient, consider a sequence $(a_i; i=1, 2, \cdots)$ of elements of $L(\mu)$ with a sum $a \in L$. Define $c_i \equiv \sum(a_j; j=1, 2, \cdots, i)$. Clearly $a=\sum c_i$, and $(c_i; i=1, 2, \cdots)$ is increasing. By Theorem 1, $c_i \in L(\mu)$ for each $i=1, 2, \cdots$, and hence $\mu(c_i+b)+\mu(c_ib)=\mu(c_i)+\mu(b)$ for each $b \in L$. On taking the limit and using B^+ we see that $\mu(a+b)+\mu(ab)=\mu(a)+\mu(b)$. Thus $a \in L(\mu)$, and B^+ is sufficient. For the necessity, consider an increasing sequence $(a_i; i=1, 2, \cdots)$ of elements of $L(\mu)$ with a sum $a \in L$. For each $b \in L$ and each $i=1, 2, \cdots$, $a+b \geqq a_i+b$ and $ab \geqq a_ib$; and hence $\mu(a+b) \geqq \mu(a_i+b)$, $\mu(ab) \geqq \mu(a_ib)$. Define $\alpha \equiv \lim \mu(a_i+b)$ and $\beta \equiv \lim \mu(a_ib)$ as $i\to\infty$. Since $a_i \in L(\mu)$ we have $\mu(a_i+b)+\mu(a_ib)=\mu(a_i)+\mu(b)$. On taking the limit and using $B^+(\mu)$ and the fact that $a \in L(\mu)$ we obtain $\alpha+\beta=\mu(a)+\mu(b)$. It follows that $\alpha=\mu(a+b)$, $\beta=\mu(ab)$. Thus B^+ is necessary when L satisfies $B^+(\mu)$. The alternate reading is dual. The proof is complete.

DEFINITION 2. (1) *For each* $a \in L$ *we define* $\mu^+(a) \equiv$ g.l.b. $[\mu(c); c \in L(\mu), c \geqq a]$, $\mu^-(a) \equiv$ l.u.b. $[\mu(c); c \in L(\mu), c \leqq a]$.

[7] Cf. L. R. Wilcox and the author, op. cit., p. 317.

(2) *We say that* $\mu(a)$ *is outer (inner) regular*[8] *in case* $\mu(a)=\mu^+(a)$ $(\mu(a)=\mu^-(a))$ *for every* $a \in L$.

LEMMA 1. *If* $\mu(a)$ *is outer regular, then*

$$\mu(a+b)+\mu(ab) \leqq \mu(a)+\mu(b)$$

for every $a, b \in L$.

PROOF. Consider $a, b \in L$. For each $c, d \in L(\mu)$ for which $c \geqq a$, $d \geqq b$ we have $c+d \geqq a+b$, $cd \geqq ab$; and, by Theorem 1, $c+d$, $cd \in L(\mu)$. Consequently, since $\mu(a)$ is outer regular, $\mu(a+b)+\mu(ab) \leqq \mu(c+d)+\mu(cd)=\mu(c)+\mu(d)$. The lemma follows by applying a simple property of the greatest lower bound.

THEOREM 3. *If* L *satisfies* $B^+(\mu)$ $(B^-(\mu))$ *and* $\mu(a)$ *is outer (inner) regular, then* L *satisfies* B^+ (B^-).

PROOF. This follows from Lemma 1 and its dual by the method used in proving Theorem 2.

We now assume that L is closed with respect to denumerable sums and products.

LEMMA 2. *If* L *satisfies* B^- (B^+), *then for each* $a \in L$ *there is an element* $c \in L(\mu)$ *such that* $c \geqq a$ $(c \leqq a)$ *and* $\mu(c)=\mu^+(a)$ $(\mu(c)=\mu^-(a))$.

PROOF. This is an easy consequence of Theorem 2.

REMARK 3. *It is now clear that when* L *satisfies* B^- *and* $\mu(a)$ *is outer regular the distance function*[9] $\delta(a, b)=2\mu(a+b)-\mu(a)-\mu(b)$ *identifies each* $a \in L$ *with an element* $c \in L(\mu)$.

THEOREM 4. *If* L *satisfies* B^+ (B^-) *and* $\mu(a)$ *is outer (inner) regular, then an element* $a \in L$ *belongs to* $L(\mu)$ *if and only if* $\mu^-(a)=\mu(a)$ $(\mu^+(a)=\mu(a))$.

PROOF. Consider an element $a \in L$ for which $\mu^-(a)=\mu(a)$. By hypothesis and Lemma 2 there is an element $c \in L(\mu)$ such that $c \leqq a$ and $\mu(c)=\mu^-(a)$. Thus, for each $b \in L$, $\mu(a)+\mu(b)=\mu^-(a)+\mu(b)=\mu(c)+\mu(b)=\mu(c+b)+\mu(cb) \leqq \mu(a+b)+\mu(ab)$. Consequently, by Lemma 1, $a \in L(\mu)$. The converse is trivial. The alternate reading is dual. The proof is complete.

LEHIGH UNIVERSITY

[8] Cf. Carathéodory, op. cit., p. 258.

[9] See L. R. Wilcox and the author, op. cit., p. 311.

THE INEQUALITIES OF MORSE WHEN THE MAXIMUM TYPE IS AT MOST THREE[1]

ROSS H. BARDELL

1. **Introduction.** The theory of critical points has been developed by Morse[2] [1, 2] by the use of combinatorial topology. In particular the theory developed by Morse is applicable to simple integral problems in the calculus of variations. The author, in his doctoral dissertation [3], studied the theory of types of extremals for a simple integral problem in the plane without the use of combinatorial topology. In the present paper a new and interesting property of extremal arcs joining two fixed points in the plane will be proved. This property makes possible a proof of the inequalities of Morse [3, p. 30] in the special case where the maximum type of each extremal arc is at most three. This is the first proof of the inequalities without the use of topology and without assuming the problem to be reversible [3, p. 25].

The hypotheses of this paper are those made in §1 of the dissertation [3] referred to above and likewise the notation used here is that of the earlier paper.

2. **Properties of the extremal arcs joining two fixed points of R.** In the paper mentioned above [3, p. 17] it was proved that every point 2 of R, which is not on an envelope arc of the family of extremal arcs though the point 1, is joined to 1 by $2r+1$ (r a positive integer) extremal arcs of which r are of odd type and $r+1$ are of even type, one of which at least is of type zero.

Consider now a point 2 of R which is not on an envelope arc of the family of extremals through the point 1. Let E_{a_0} designate an extremal arc joining the point 2 to 1 and such that the arc 12 of E_{a_0} is of type zero. It can be shown [3, p. 10] that the extremal arc E_a for $a>a_0$ and near a_0 and also for $a<a_0+2\pi$ and near $a_0+2\pi$ has no intersections with the arc 12 of E_{a_0}. Then as a increases from a_0 to $a_0+2\pi$ all intersections of the extremal arc E_a with the arc 12 of E_{a_0} which move onto this arc must move off again. In fact it can easily be proved [3, p. 18] that an intersection 3 of E_a with the arc 12 of E_{a_0} which moves onto the arc 12 of E_{a_0} for an extremal arc of odd (even) type must move off of the arc 12 of E_{a_0} for an extremal arc of even

[1] Presented to the Society, April 8, 1938.

[2] The numbers in brackets here and elsewhere refer to the bibliography at the end of this paper.

(odd) type. While the preceding statement is true for all extremal arcs, the following much stronger property can be proved in the special case where the extremal arcs in question are of type zero or one.

LEMMA 2.1. *Let 2 be any point of the region R which is not on an envelope arc of the family of extremals through 1, and let 3 designate a particular intersection of E_a with the arc 12 of E_{a_0} which is moving through 2 onto the arc 12 of E_{a_0} as a increases through a' and off of the arc 12 of E_{a_0} as a increases through a'', where $a'<a''$. Then, if the arc 12 of $E_{a'}$ is of type zero, the arc 12 of $E_{a''}$ is of type one. Moreover, if the arc 12 of $E_{a''}$ is of type zero, then the arc 12 of $E_{a'}$ is of type one.*

It will suffice to prove the first statement of the above lemma since the last statement is exactly equivalent to the first when the parameter is allowed to decrease from a'' to a'. We will now proceed to prove the first statement.

Since, by hypotheses, the arc 12 of $E_{a'}$ is of type zero, it follows that the arc 13 of E_a for $a>a'$ and near a' has no intersections with the arc 12 of $E_{a'}$ [3, p. 10]. Moreover, as a increases from a' to a'' the point 3 never passes through 2, and hence no intersection of the arc 13 of E_a can move onto the arc 12 of $E_{a'}$. Therefore the arc 13 of E_a for $a'<a<a''$ has no intersections with the arc 12 of $E_{a'}$ and hence the arc 12 of $E_{a''}$ has only the point 2 in common with the arc 12 of $E_{a'}$.

Let the arc 12 of $E_{a''}$ be of type k. Then as a decreases from a'' and is near a'', the arc 13 of E_a will have k intersections with the arc 12 of $E_{a''}$. We wish to show that the number of intersections of the arc 13 of E_a with the arc 12 of $E_{a''}$ is also equal to k for $a \leqq a'$ and near a'. In that case the preceding paragraph shows that $k=1$ and hence the lemma is proved. To show this let 4 and 6 designate the intersections of $E_{a'}$ and $E_{a''}$, respectively, with E_{a_0} which are adjacent to and preceding 2 on E_{a_0} and let 5 designate the intersection of E_a with E_{a_0} which is adjacent to and preceding 3 on E_{a_0}. Further let the intersection of $E_{a''}$ with E_{a_0} which is adjacent to and preceding 6 on E_{a_0} be designated by 7. We have then the following two cases to consider.

CASE 1. Sin $(\theta_0-\theta)$ has opposite signs at the two points 4 and 2, where θ_0 and θ are the inclinations of the tangents to E_{a_0} and E_a, respectively, at their particular intersection in question. In this case the intersections 5 and 3 can move only on the interval 42 of E_{a_0} for $a'<a<a''$. This follows since, as was seen above, the arc 13 of E_a has no intersections with the arc 12 of $E_{a'}$, for $a'<a<a''$. Therefore the intersection 6 must also be on the interval 42 of E_{a_0} and by exactly

the same reasoning it follows that the intersection 7 must be on the interval 14 of E_{a_0}. Hence as a decreases from a'' to a' the intersection 3, which must stay on the interval 42, may pass through 6 but not through 7. When it does pass through 6 an intersection of the arc 13 of E_a with the arc 12 of $E_{a''}$ is lost, but is regained when 3 reverses its direction and moves again through 6 as it must since its final position is 2. Therefore in this case the arc 13 of E_a for $a \leqq a'$ and near a' has k intersections with the arc 12 of $E_{a''}$.

CASE 2. Sin $(\theta_0 - \theta)$ has the same sign at the two intersections 4 and 2. In this case the point 3 can move only on the interval 42 of E_{a_0} and the point 5 can move only on the interval 14 of E_{a_0} as a decreases from a'' to a'. This follows from the fact that, as above, the arc 13 of E_a has no intersections with the arc 12 of $E_{a'}$ for $a' < a < a''$. Therefore 6 must remain on the interval 14 of E_{a_0} as a decreases from a'' to a' and hence 3 cannot pass through 6 since 3 must remain on the interval 42 of E_{a_0}. Hence in this case also the arc 13 of E_a for $a' \leqq a < a''$ always has k intersections with the arc 12 of $E_{a''}$.

3. **The inequalities.** For every a on the interval $a_0 < a < a_0 + 2\pi$ defining an extremal arc E_a which joins the point 2 to 1 an intersection of E_a with the arc 12 of E_{a_0} either moves onto or off of this arc. Let the extremal arcs joining the points 1 and 2 be paired so that the arcs of each pair are those defined in the family of extremals through the point 1 by the parameter values for which a particular intersection 3 of E_a with the arc 12 of E_{a_0} moves onto and off of this latter arc.

LEMMA 3.1. *Every extremal arc of the set of extremals joining the points* 1 *and* 2, *except* E_{a_0}, *belongs to one and only one of the pairs of extremal arcs defined above. Moreover, if the maximum type of each extremal arc joining the points* 1 *and* 2 *is assumed to be not greater than* 3, *then the extremal arcs of each pair differ in type by unity.*

The first statement of the lemma is an immediate consequence of the definition of the pairs of extremal arcs given in the first paragraph of this section. The last statement in the lemma is a consequence of Lemma 2.1 and the statement in the paragraph immediately preceding this lemma.

Let p_i represent the number of pairs of extremal arcs joining the points 1 and 2, such that one arc of each pair is of type i and the other of type $i-1$. Obviously $p_i \geqq 0$, $(i = 1, 2, 3)$. Let M_i represent the number of extremal arcs of type i joining the points 1 and 2. Then equations (4:1) and (4:2), on page 24 of the dissertation [3] referred to above, with $S_i = M_i$ are easily obtained. Moreover, the in-

equalities (4:3) on the same page, with $S_i = M_i$, $(i = 0, 1, 2, 3)$, follow as before except that the $>$ must be replaced by $\geqq$ since now $p_i \geqq 0$.

Summarizing the results of the preceding pages we have the following theorem.

THEOREM 3.1. *For a calculus of variations problem and a region R satisfying the hypotheses made in §1 above, let 1 and 2 represent two fixed points of R such that 2 is not an envelope arc of the family of extremals through 1. Further let M_i represent the number of extremal arcs joining the points 1 and 2 and each of type i and let 3 be the maximum type of each arc. Then the numbers M_i, $(i = 0, 1, 2, 3)$, so defined must satisfy the relations* $M_0 \geqq 1$, $M_1 \geqq M_0 - 1$, $M_2 \geqq M_1 - M_0 + 1$, $M_3 = M_2 - M_1 + M_0 - 1$.

BIBLIOGRAPHY

1. M. Morse, *The Calculus of Variation in the Large*, American Mathematical Society Colloquium Publications, vol. 18, New York, 1934.

2. M. Morse, *Functional topology and abstract variational theory*, Annals of Mathematics, (2), vol. 38 (1937), pp. 386–449.

3. R. H. Bardell, *Contributions to the Calculus of Variations*, 1933–1937, University of Chicago Press, 1937, pp. 279–312.

UNIVERSITY OF WISCONSIN AT MILWAUKEE

A METHOD FOR PROVING CERTAIN ABSTRACT GROUPS TO BE INFINITE[1]

H. S. M. COXETER

1. **Introduction.** I have stated elsewhere[2] that the group $(3, 3, 4; 4)$, defined by

$$R^3 = S^3 = (RS)^4 = (R^{-1}S^{-1}RS)^4 = 1,$$

is infinite. This fact will now be proved by showing that there is a factor group of order $24n^4$ for every positive integer n.

We shall find a closely related group of order $48n^4$, satisfying the relations $S^2 = T^2 = (ST)^8 = (S^{-1}TST)^6 = 1$, which have been studied by Brahana;[3] but there is no overlapping, since his "subgroup H" is not invariant in our case, although there still is an abelian invariant subgroup of index 48. In fact, it was the search for such a subgroup that led to the simple treatment given here.

Section 7 is inserted for its intrinsic interest, and can be omitted without impairing the proof of the main result (§8).

2. **A group of order n^4.** Consider the direct product of two cyclic groups of order n. Since the defining relations $M_1^n = M_2^n = M_1^{-1}M_2^{-1}M_1M_2 = 1$ imply $(M_1M_2)^n = 1$, they may be put into the form[4]

$$M_1^n = M_2^n = M_3^n = M_1M_2M_3 = M_3M_2M_1 = 1. \tag{1}$$

Hence the direct product of four cyclic groups of order n is defined by

$$\begin{aligned} M_i^n = M_1M_2M_3 = M_3M_2M_1 = N_j^n = N_1N_2N_3 = N_3N_2N_1 = 1, \\ M_iN_j = N_jM_i, \qquad i, j = 1, 2, 3. \end{aligned} \tag{2}$$

3. **A group of order $4n^4$.** These relations continue to hold when M_i is replaced by N_i, and N_j by M_j^{-1}. We now enlarge the group of order n^4 by adjoining an operator A, of period four, which transforms it according to this automorphism. The extra relations that have to be added to (2) are

$$A^4 = 1, \qquad A^{-1}M_iA = N_i, \qquad A^{-1}N_jA = M_j^{-1}.$$

---

[1] Presented to the Society, September 6, 1938. The enumerative method described in the abstract (this Bulletin, 44-9-331) seems to be effective only in those cases where more orthodox methods are equally effective.

[2] Coxeter [2, p. 101, second footnote].

[3] Brahana [1].

[4] In the notation of Coxeter [2, p. 87], this is $(n, n, n; 1)$.

The enlarged group,[5] of order $4n^4$, may be put into a symmetrical form by defining $B=AM_3$, $C=N_2A$, and eliminating the M's and N's by means of the relations

$$B^{-1}C = M_3^{-1}M_2^{-1} = M_1, \quad C^{-1}A = A^{-1}N_2^{-1}A = M_2, \quad A^{-1}B = M_3,$$
$$BC^{-1} = N_3^{-1}N_2^{-1} = N_1, \quad CA^{-1} = N_2, \quad AB^{-1} = AM_3^{-1}A^{-1} = N_3.$$

The result is

$$(3) \quad \begin{aligned} A^4 = B^4 = C^4 = (BC)^2 = (CA)^2 = (AB)^2 = A^{-1}BC^{-1}AB^{-1}C \\ = (B^{-1}C)^n = (C^{-1}A)^n = (A^{-1}B)^n = 1. \end{aligned}$$

These relations imply

$$\begin{aligned} (A^2B^{-1}C)^2 = (B^2C^{-1}A)^2 = (C^2A^{-1}B)^2 = (A^{-1}B^2C)^2 = (B^{-1}C^2A)^2 \\ = (C^{-1}A^2B)^2 = (A^2B^2C^2)^2 = (ABC)^4 = 1. \end{aligned}$$

In detail,

$$\begin{aligned} (A^2B^{-1}C)^2 &= A^{-1}\cdot A^{-1}B^{-1}\cdot CA\cdot AB^{-1}C = A^{-1}\cdot BA\cdot A^{-1}C^{-1}\cdot AB^{-1}C \\ &= A^{-1}BC^{-1}AB^{-1}C = 1, \\ (A^{-1}B^2C)^2 &= A^{-1}B\cdot BC\cdot A^{-1}B^{-1}\cdot B^{-1}C = A^{-1}B\cdot C^{-1}B^{-1}\cdot BA\cdot B^{-1}C \\ &= A^{-1}BC^{-1}AB^{-1}C = 1, \\ (A^2B^2C^2)^2 &= A^2B\cdot BC\cdot CA^2B^2C^2 = A^2BC^{-1}\cdot B^{-1}CA^2\cdot B^2C^2 \\ &= CB^{-1}A^{-2}\cdot A^{-2}C^{-1}B\cdot B^2C^2 = C\cdot B^{-1}C^{-1}B^{-1}\cdot C^2 = C^4 = 1, \\ (ABC)^4 &= (AB\cdot CA\cdot BC)^2 = (B^{-1}A^{-1}\cdot A^{-1}C^{-1}\cdot C^{-1}B^{-1})^2 \\ &= (BC^2A^2B)^{-2} = 1. \end{aligned}$$

As one of the relations $M_i^n=1$ is superfluous in (1), so one of the consequent relations $(B^{-1}C)^n=(C^{-1}A)^n=(A^{-1}B)^n=1$ is superfluous in (3), say $(C^{-1}A)^n=1$. In terms of A, B, C and $(ABC)^{-1}$, (3) takes the form

$$(4) \quad \begin{aligned} A^4=B^4=C^4=D^4 = ABCD = (BC)^2 = (CA)^2 = (AB)^2 = (BD)^2 \\ = (A^{-1}B)^n = (B^{-1}C)^n = (C^{-1}D)^n = (D^{-1}A)^n = 1, \end{aligned}$$

implying

$$BC^{-1}AB^{-1}CA^{-1} = BC^{-1}\cdot AB\cdot B\cdot BC\cdot A^{-1} = (BC^{-1}B^{-1}A^{-1})^2 = (BD)^2 = 1.$$

Of course, the relations $(C^{-1}D)^n=(D^{-1}A)^n=1$ (inserted for the

[5] There is an intermediate group, of order $2n^4$, generated by A^2, AB, B^2, BC, C^2, CA, and defined by $T_1^2=T_2^2=T_3^2=T_4^2=T_5^2=T_6^2=T_1T_2T_3T_4T_5T_6=(T_1T_3)^n=(T_2T_3)^n=(T_3T_4)^n=(T_4T_5)^n=(T_5T_6)^n=(T_6T_1)^n=(T_iT_jT_k)^2=1$, $(i<j<k)$.

sake of symmetry) are superfluous; in fact, the rest of (4) implies $AB^{-1}\cdot A^{-1}C=A^{-1}C\cdot AB^{-1}$, whence $(AD^{-1})^n=(A^2BC)^n=(AB^{-1}\cdot A^{-1}C)^n$ $=(AB^{-1})^n(A^{-1}C)^n=1$, and similarly $(D^{-1}C)^n=(ABC^2)^n=(AC^{-1}$ $\cdot B^{-1}C)^n=(AC^{-1})^n(B^{-1}C)^n=1$.

4. **A group of order $8n^4$.** The relations (4) continue to hold when A, B, C and D are replaced by C^{-1}, D^{-1}, D^2A $(=DA^{-1}D^{-1})$ and BC^2 $(=C^{-1}B^{-1}C)$. For, BC, CA, AB, BD, $A^{-1}B$, $B^{-1}C$, $C^{-1}D$, $D^{-1}A$ are then replaced by DA, DB, DC, AC, CD^{-1}, $D^{-1}A$, BA^{-1}, $C^{-1}B$. Let $\mathcal{P}$ denote this automorphism. The repeated automorphism $\mathcal{P}^2$ replaces A, B, C, D by DAD^{-1}, $C^{-1}BC$, BCB^{-1}, $A^{-1}DA$, and so is equivalent to transformation by BC or DA. We now enlarge the group of order $4n^4$ by adjoining an operator P, whose square is BC, and which transforms the group according to the automorphism $\mathcal{P}$. The extra relations that have to be added to (4) are

$$P^2 = BC,\ P^{-1}AP = C^{-1},\ P^{-1}BP = D^{-1},\ P^{-1}CP = D^2A,\ P^{-1}DP = BC^2.$$

Defining $Q=AP$, so that $QP^{-1}=A$, $PQ=B$, $Q^{-1}P=C$, $P^{-1}Q^{-1}=D$, we obtain the enlarged group, of order $4n^4$, in the form

$$\begin{aligned}(5)\qquad P^4 = Q^4 = (PQ)^4 = (P^{-1}Q)^4 = (P^{-1}Q^{-1}PQ)^2\\ = (PQ^{-1}PQ)^n = (P^{-1}QPQ)^n = 1.\end{aligned}$$

Since

$$\begin{aligned}(PQ^2)^2 &= PQ\cdot QP\cdot Q^2 = BD^{-1}AB = BAB\cdot CA\cdot B\\ &= A^{-1}\cdot A^{-1}C^{-1}\cdot B = (B^{-1}CA^2)^{-1},\\ (P^2Q)^2 &= P^2\cdot QP\cdot PQ = BCD^{-1}B = B\cdot CA\cdot BCB\\ &= B\cdot A^{-1}C^{-1}\cdot C^{-1} = (C^2AB^{-1})^{-1},\\ P^{-1}Q^2P^{-1} &= C^{-1}A,\end{aligned}$$

the relations (5) must imply $(PQ^2)^4=(P^2Q)^4=(P^2Q^2)^n=1$. In terms of P, Q and $(PQ)^{-1}$, we have, therefore,

$$\begin{aligned}(6)\qquad O^4 = P^4 = Q^4 = OPQ = (QPO)^2 = (P^{-1}Q)^4 = (Q^{-1}O)^4\\ = (O^{-1}P)^4 = (P^2Q^2)^n = (Q^2O^2)^n = (O^2P^2)^n = 1.\end{aligned}$$

5. **A group of order $24n^4$.** To this group of order $8n^4$ we adjoin an operator R, of period three, which transforms the three generators according to a cyclic permutation.[6] The substitution $O=RQR^{-1}$, $P=R^{-1}QR$ gives us the enlarged group, of order $24n^4$, in the form

[6] Compare Coxeter [2, p. 96].

(7) $Q^4 = R^3 = (QR)^3 = (Q^{-1}R)^6 = (Q^{-1}R^{-1}QR)^4 = (Q^2R^{-1}Q^2R)^n = 1.$

In terms of R and $(QR)^{-1}$, this becomes

(8) $R^3 = S^3 = (RS)^4 = (R^{-1}S)^6 = (R^{-1}S^{-1}RS)^4 = (R^{-1}SRS^{-1}RS)^n = 1.$

6. **A group of order $48n^4$.** Finally, to the group of order $24n^4$ we adjoin an involutory operator T which interchanges R and S, obtaining

$$(9) \qquad \begin{aligned} S^3 = T^2 = (ST)^8 = (S^{-1}TST)^6 &= [(S^{-1}T)^2(ST)^2]^4 \\ &= [(ST)^4T]^{2n} = 1. \end{aligned}$$

For, if $T^2=1$ and $R=TST$, we have $R^{-1}S^{-1}RS=TS^{-1}TS^{-1}TSTS$ and $R^{-1}SRS^{-1}RS=(TS^{-1}TSTS)^2=(TST\cdot T\cdot STSTS)^2$.

In terms of ST and T, (9) takes the form

(10) $U^8 = T^2 = (UT)^3 = (U^{-1}TUT)^6 = (U^{-2}TU^2T)^4 = (U^4T)^{2n} = 1.$

7. **Other related groups.** Several further groups with simple defining relations can be derived from those obtained above. For instance, adjoining to (3) an operator V which cyclically permutes A, B, C, we obtain the group

(11) $V^3 = A^4 = (V^{-1}AVA)^2 = (VA^{-1}VA)^3 = (V^{-1}A^{-1}VA)^n = 1,$

of order $12n^4$, and we deduce that these relations imply $(VA)^{12} = (VA^2)^6 = 1$.

Again, adjoining to (4) an operator X which cyclically permutes A, B, C, D, we obtain

$$X^4 = A^4 = (XA)^4 = (X^2A)^4 = (X^{-1}AXA)^2 = (X^{-1}A^{-1}XA)^n = 1,$$

of order $16n^4$. In terms of X and XA, this becomes

(12) $X^4 = Y^4 = (XY)^4 = (X^{-1}Y)^4 = (X^2Y^2)^2 = (X^{-1}Y^{-1}XY)^n = 1.$

Concerning (5), it is natural to ask whether the periods of $PQ^{-1}PQ$ and $P^{-1}QPQ$ are inevitably equal. The rather surprising answer is, as we shall see, that by leaving one of them unrestricted we only double the order of the group. Since O, P, Q are interchangeable, this means that the group

(13) $P^4 = Q^4 = (PQ)^4 = (P^{-1}Q)^4 = (P^{-1}Q^{-1}PQ)^2 = (P^2Q^2)^n = 1$

is of order $16n^4$, like (12).

To build up such a group, we begin with the direct product of two cyclic groups of orders $2n$ and n (generated by M_2 and M_3), which can be written in a form resembling (1):

$$M_1^n = M_2^n, \qquad M_3^n = M_1M_2M_3 = M_3M_2M_1 = 1.$$

Instead of (2) we take the group

$$M_1^n = M_2^n = N_1^n = N_2^n,$$
$$M_3^n = M_1M_2M_3 = M_3M_2M_1 = N_3^n = N_1N_2N_3 = N_3N_2N_1 = 1,$$
$$M_iN_j = N_jM_i, \qquad i, j = 1, 2, 3,$$

whose order is $2n^4$ since its general operator can be expressed as

$$M_1^pM_2^qN_1^rN_2^sZ^t, \qquad 0 \leqq p, q, r, s < n; t = 0 \text{ or } 1,$$

where $Z = M_1^n$.

Instead of (3) and (4), we derive two equivalent definitions for a certain group of order $8n^4$: first[7]

$$A^4 = B^4 = C^4 = (BC)^2 = (CA)^2 = (AB)^2 = A^{-1}BC^{-1}AB^{-1}C$$
$$= (AC^{-1})^n(B^{-1}C)^n = (A^{-1}B)^n = 1,$$

and second

$$A^4 = B^4 = C^4 = D^4 = ABCD = (BC)^2 = (CA)^2 = (AB)^2 = (BD)^2$$
$$= (A^{-1}B)^n = (C^{-1}D)^n = 1.$$

Finally, instead of (5) we obtain the group, of order $16n^4$,

$$P^4 = Q^4 = (PQ)^4 = (P^{-1}Q)^4 = (P^{-1}Q^{-1}PQ)^2 = (PQ^{-1}PQ)^n = 1.$$

8. **Conclusions regarding infinite groups.** The consistency of (8) for all values of n shows that the group $(3, 3|4, 6; 4)$, defined by $R^3 = S^3 = (RS)^4 = (R^{-1}S)^6 = (R^{-1}S^{-1}RS)^4 = 1$, is infinite. The "larger" groups[8] $(3, 3|4, 6)$, $(3, 3, 4; 4)$ are infinite *a fortiori*. Similarly, (5) establishes infinite order for $(4, 4|4, 4; 2)$, and thence for[9] $(4, 4, 4; 2)$.

9. **Comparison with Brahana's groups.** The infinite group $(2, 3, 8; 6)$, of which (9) is a factor group, has been investigated by Brahana.[10] His operators T_1, T_2, T_3 are easily recognized in our factor group as BC, AB, D^2. Since $T_2T_3T_1 = CD \cdot D^2 \cdot DA = CA$, the subgroup

[7] These relations imply $(B^{-1}C)^n = (C^{-1}A)^n = (BC^{-1})^n = (CA^{-1})^n$. In detail, $(C^{-1}A)^{-n}(B^{-1}C)^n = (C^{-1}A)^{-n}(CA^{-1})^n = A^{-1}(AC^{-1})^{-n}(ACA^2)^nA = A^{-1}(B^{-1}C)^n(C^{-1}A)^nA = A^{-1}(B^{-1}A)^nA = 1$.

[8] Coxeter [2, pp. 86, 101].

[9] Coxeter [2, p. 97]. By the method of Coxeter [2, p. 90, §2.5], (12) establishes infinite order for $(4, 8|2, 4; 4)$. This raises an interesting question as to the finite or infinite order of $(4, 7|2, 4; 4)$.

[10] Brahana [1, p. 892].

$\{T_1, T_2, T_3\}$ is $\{BC, CA, AB\}$. This subgroup, being[11] $((n, n, n; 2))$, of order $2n^2$, is of index $24n^2$. It is not invariant,[12] since, if it were, its index would be just 24. Hence (9) is not one of the groups treated in Brahana's main theorems, but is a first step towards the "large undertaking" mentioned in his final paragraph.[13]

References

1. H. R. Brahana, *On the groups generated by two operators of orders two and three whose product is of order eight*, American Journal of Mathematics, vol. 53 (1931), pp. 891–901.

2. H. S. M. Coxeter, *The abstract groups* $G^{m,n,p}$, Transactions of this Society, vol. 45 (1939), pp. 73–150.

University of Toronto

[11] Coxeter [2, p. 143].
[12] Brahana [1, p. 892].
[13] Brahana [1, p. 901].

ANOMALOUS PLANE CURVE SYSTEMS ASSOCIATED WITH SINGULAR SURFACES[1]

T. R. HOLLCROFT

1. **Introduction.** A plane curve system of order n and genus p, with δ nodes, κ cusps, ι statangents, τ bitangents, is of first virtual dimension d_0, of virtual dimension d_0', and of effective dimension d. For such a system, $d_0' = 3n+p-\kappa-1 = 3m+p-\iota-1$. If the curve system has only distinct nodes and cusps, $d_0 = d_0'$; if higher singularities, $d_0 < d_0'$.

For any irreducible, continuous, complete curve system, $d \geqq d_0$. If $d = d_0$, the curve system is called homalous, if $d > d_0$, anomalous. The anomaly A of a plane curve system is defined by the relation $A = d - d_0$. The above definitions were introduced by B. Segre.[2]

The sections by a plane π of the tangent cones to a nonsingular surface of order ν from the points of S_3 constitute a continuous plane curve system with the characteristics:

$$n = \nu(\nu-1), \qquad \kappa = \nu(\nu-1)(\nu-2),$$
$$\delta = (1/2)\nu(\nu-1)(\nu-2)(\nu-3).$$

B. Segre[3] has shown that for this curve system,

$$d_0 = (5/2)\nu(\nu-1), \qquad d = (1/6)(\nu+1)(\nu+2)(\nu+3) - 5,$$
$$A = d - d_0 = (1/6)(\nu-2)(\nu-3)(\nu-4).$$

Hence, for $\nu \geqq 5$, this plane curve system is anomalous.

The purpose of the present paper is to ascertain the dimensions and anomaly of plane sections of tangent cones to certain singular surfaces.

---

[1] Presented to the Society, April 16, 1938.

[2] B. Segre, *Esistenza e dimensione di sistemi continui distinti di curve piane algebriche con dati caratteri*, Rendiconti dell'Accademia dei Lincei, (6), vol. 10 (1929), pp. 31–38. The adjective "irregolare" used by B. Segre to describe the plane curve system for which $d > d_0$ has two English equivalents "irregular" and "anomalous." The first translation is the more natural one. Since it has been found, however, that similar systems of surfaces exist and since the term "irregular" has long had a different definite meaning for surfaces, another term is necessary for surfaces and, therefore, should also be used for plane curves. Homalous and anomalous are Greek antonyms. Only the latter has heretofore been used in English.

[3] B. Segre, *Sulla caratterizzazione delle curve di diramazione dei piani multipli generali*, Memorie delle Reale Accademia d'Italia, Classe di Scienze Fisische, Matematiche e Naturali, vol. 1 (1930), pp. 5–31.

2. **Surfaces with multiple points.** Let the surface F of order ν have the point P of multiplicity k, such that the plane section of the tangent cone to F at P is a nonsingular curve of order k. The tangent cone to F from an arbitrary point has a multiple generator of order $k(k-1)$ through P. This multiple generator reduces the number of distinct cuspidal generators by $k(k-1)(k-2)$ and the number of distinct nodal generators by $(1/2)k(k-1)(k-2)(k-3)$.

The plane sections C by a plane π of the tangent cones to F from points of S_3 consist of plane curves each of which has a $k(k-1)$-fold point at the intersection of π with the multiple generator. This plane curve system C has the additional characteristics

$$n = \nu(\nu-1), \qquad m = \nu(\nu-1)^2 - k(k-1)^2,$$

$$p = (1/2)\nu(\nu-1)(2\nu-5) - (1/2)k(k-1)(2k-3),$$

$$\kappa = \nu(\nu-1)(\nu-2) - k(k-1)(k-2),$$

$$\delta = (1/2)\nu(\nu-1)(\nu-2)(\nu-3) - (1/2)k(k-1)(k-2)(k-3),$$

$$d_0 = (5/2)\nu(\nu-1) - (3/2)k(k-1) + 2.$$

In the above, δ and κ are the numbers of distinct nodes and cusps of the plane sections. The $k(k-1)$-fold point contains also the equivalent of $(1/2)k(k-1)(k^2-k-1)$ nodes.

To obtain the effective dimension of C, the effective dimension of F must be found. The invariant postulation[4] of a k-fold point on F is $(1/6)k(k+1)(k+2)-3$. Since the above plane curve system may be the plane section of the tangent cone to ∞^4 surfaces, the effective dimension d of C is

$$d = (1/6)(\nu+1)(\nu+2)(\nu+3) - (1/6)k(k+1)(k+2) - 2.$$

Hence

$$A = (1/6)(\nu-2)(\nu-3)(\nu-4) - (1/6)(k-1)(k-2)(k-3).$$

The above formulas are derived for one multiple point of order k. If the surface has α distinct, independent, multiple points of orders k_i, the above formulas hold with the sum sign. For example, for such a surface,

$$A = (1/6)(\nu-2)(\nu-3)(\nu-4) - (1/6)\sum_{i=1}^{\alpha}(k_i-1)(k_i-2)(k_i-3).$$

From the above expression for A, there results:

(1) A surface may have any number of independent double or

[4] T. R. Hollcroft, *Invariant postulation*, this Bulletin, vol. 36 (1930), p. 421.

triple points without affecting the anomaly of plane sections of the tangent cone.

(2) The plane sections of the tangent cones to a monoid ($\alpha=1$, $k=\nu-1$) are always homalous plane curve systems.

Since $A\not\equiv 0$, this expression for A ($k_i=k$) defines the following limit to the number α of distinct independent k-fold points, $k\geqq 4$, on an irreducible surface of order ν:

$$\text{(A)} \qquad \alpha \leqq \frac{(\nu-2)(\nu-3)(\nu-4)}{(k-1)(k-2)(k-3)}.$$

The smallest limit obtained heretofore results from the fact that the genus of the plane section of the tangent cone to an irreducible surface cannot be negative. This limit[5] is

$$\text{(B)} \qquad \alpha \leqq \frac{(\nu-2)(2\nu^2-3\nu-1)}{k(k-1)(2k-3)}.$$

Limit (A) is smaller than limit (B) for a given value of $k\geqq 4$ and $\nu=k+\beta$, where

$$\beta \leqq \frac{3(6k-k^2-6)+(121k^4-832k^3+1988k^2-1920k+612)^{1/2}}{2(7k-12)}.$$

A very close approximation to this limiting value of β is given by $\beta\leqq(1/7)(4k-3)$.

It has been assumed that the multiple points are unrelated in position on the surface and that they are associated with sets of independent invariants among the coefficients of the surface.

The nodes and cusps of the system C lie on adjoint curves. The adjoint curves of the lowest order are the following:

(1) The adjoint curves of order $(\nu-1)(\nu-2)$ pass simply through the $\delta+\kappa$ distinct double points and have a multiple point of order $(k_i-1)(k_i-2)$ at each $k_i(k_i-1)$-fold point.

(2) The adjoint curves of order $\nu(\nu-2)$ pass simply through the nodes, are tangent to C at each cusp, and have a multiple point of order $k_i(k_i-2)$ at each $k_i(k_i-1)$-fold point.

If the multiple points have cones whose plane sections are singular curves, the values of d and d_0 can be found in a given case, but a general formula that will include all such cases is not feasible.

---

[5] T. R. Hollcroft, *Limits for multiple points and curves of surfaces*, Tôhoku Mathematical Journal, vol. 30 (1928), pp. 116–117. Limit (1) derived on page 116 depends upon the Lefschetz postulate and, therefore, does not hold.

3. **Surfaces with singular curves.** Consider the surface F of order ν and class ν' with a nodal curve of order b and a cuspidal curve of order c. The other surface characteristics to be used are the following:

a, order of tangent cone = class of plane section;
δ, number of nodal generators of tangent cone;
κ, number of cuspidal generators of tangent cone;
σ, class of cuspidal developable;
σ', order of parabolic curve;
j, number of pinch points on nodal curve;
ω, number of tacnodal points on cuspidal curve;
χ, number of cubic nodes on cuspidal curve;
Δ, effective dimension of surface.

From the Cayley-Zeuthen equations, there results

$$(1)\quad a = \nu(\nu - 1) - 2b - 3c, \quad 2\delta + 4\kappa = a(a - 2) + \sigma + 2j + 3\chi + 3\omega.$$

Since, for the plane sections C of the tangent cone,

$$d_0 = (1/2)a(a + 3) - \delta - 2\kappa,$$

there results, from the above,

$$(2)\qquad d_0 = (1/2)(5a - \sigma - 2j - 3\chi - 3\omega).$$

For a nonsingular surface, $d_0 = 5a/2$.

No such general formula exists for the effective dimension d of this curve system C. Further and more specific characteristics of the double curves must be given. For example, the dimension of a space curve may involve properties of the two surfaces of which it is the partial or complete intersection. In order to determine the effective dimension Δ of F, the intersection characteristics of the nodal and cuspidal curves must be known.

In particular cases, when sufficient data is given, Δ can be found, and therefore d. Nodal and cuspidal lines will be investigated.

4. **Surfaces with a nodal line.** The surface F of order ν has a nodal line both of whose tangent planes are torsal. The characteristics of F are

$$a = \nu(\nu - 1) - 2, \quad b = 1, \quad j = 2(\nu - 2), \quad \nu' = (\nu - 2)(\nu^2 - 6).$$

The virtual dimension of the plane section C of the tangent cone is $d_0 = (1/2)(5a - 2j) = (1/2)(\nu - 2)(5\nu + 1)$.

The postulation of a nodal line on F is $3\nu + 1$. Its invariant postulation[6] is $3(\nu - 1)$.

[6] *Invariant postulation*, loc. cit., p. 424.

Hence, for F

$$\Delta = (1/6)(\nu + 1)(\nu + 2)(\nu + 3) - 3\nu + 2.$$

The effective dimension of C is $d = \Delta - 4$, hence

$$d = (1/6)(\nu + 1)(\nu + 2)(\nu + 3) - 3\nu - 2.$$

The anomaly of the system C is

$$A = d - d_0 = (1/6)\nu(\nu - 4)(\nu - 5).$$

If one tangent plane of the nodal line is fixed and the other torsal, the line has $2(\nu - 2)$ pinch points that coincide in pairs. The computation is the same as above for both planes torsal.

5. **Surfaces with a cuspidal line.** A surface F may have two types of cuspidal lines, (1) the tangent plane to F is fixed, that is, is the same for all points of the line; (2) the tangent plane to F along the line is torsal.

(1) The surface F of order ν has a cuspidal line, such that the same plane is tangent to F at every point of this line. F has the characteristics

$$a = \nu(\nu - 1) - 3, \quad \nu' = (\nu - 3)(\nu^2 + \nu - 8),$$
$$\omega = \nu - 2, \quad \chi = \nu - 3, \quad \kappa = (\nu - 3)(\nu^2 - 4), \quad d_0 = (1/2)\nu(5\nu - 11),$$
$$d = \Delta - 4 = (1/6)(\nu + 1)(\nu + 2)(\nu + 3) - 4\nu - 1,$$
$$A = (1/6)\nu(\nu - 4)(\nu - 5).$$

(2) The surface F of order ν with a cuspidal line whose tangent plane is torsal has the characteristics

$$a = \nu(\nu - 1) - 3, \quad \nu' = (\nu - 3)(\nu^2 + \nu - 8),$$
$$\chi = \nu, \quad \omega = \nu - 4, \quad \sigma = 1, \quad \kappa = (\nu - 3)(\nu^2 - 4) + 4,$$
$$d_0 = (1/2)\nu(5\nu - 11) - 2,$$
$$d = \Delta - 4 = (1/6)(\nu + 1)(\nu + 2)(\nu + 3) - 4\nu - 3,$$
$$A = (1/6)\nu(\nu - 4)(\nu - 5).$$

In this case, the cuspidal developable is a point, and is, therefore, of order zero and class $\sigma = 1$.

Although the values of certain characteristics are different, the value of A is the same for both types of cuspidal lines.

6. **Summary for double lines.** The results of the two preceding sections lead to the following conclusion:

If an algebraic surface F of order $\nu \geqq 4$ has a double line (either nodal or cuspidal), the anomaly A of the plane sections C of a tangent cone to F is $A = (1/6)\nu(\nu-4)(\nu-5)$. As is evident from this expression for A, plane curve systems C associated with a surface having a double line are anomalous only for $\nu \geqq 6$.

When $\nu = 3$, the above formula yields $A = 1$. However, a cubic surface with a nodal line is ruled, and a cubic surface with a cuspidal line is a cone. The treatment given in the two preceding sections and the resulting expression for A do not apply to such surfaces.

WELLS COLLEGE

ON THE STABILITY OF THE LAMINAR FLOW OF A VISCOUS FLUID[1]

RUDOLPH E. LANGER

The problem of the effect of two-dimensional first-order disturbances upon the laminar flow of an incompressible viscous fluid is known to be fundamental to the analysis of the phenomenon of turbulence. This discussion is concerned with such a problem in the case of a flow which takes place parallel to and between parallel plane boundaries. If the direction normal to these boundaries is designed by y, and that of the flow by x, the unit of length and the origin of coordinates may be chosen so that the boundary planes are given by the equations $y = 1$ and $y = -1$. It is to be assumed then that the velocity of the undisturbed flow depends only on y, and is given by a function $U(y)$, which is suitably differentiable and nonnegative, which is non-increasing as to $|y|$, and which vanishes at the boundaries. If the maximum velocity is chosen as the unit, as will be assumed, it follows that

$$U(0) = 1, \qquad U(1) = 0, \qquad U(-y) \equiv U(y).$$

The disturbance imposed upon this flow is to be taken as an elementary wave of length $2\pi/\alpha$, in the direction of flow.

The problem as stated is known[2] to admit of formulation as the differential boundary problem

[1] Presented to the Society, September 7, 1939.

[2] C. L. Pekeris, *On the stability problem in hydrodynamics*, I, Proceedings of the Cambridge Philosophical Society, vol. 32 (1936), p. 55, and II, Journal of the Aeronautical Sciences, vol. 5 (1938), p. 236.

$$f^{iv}(y) - 2\alpha^2 f''(y) + \alpha^4 f(y)$$
$$+ i\alpha R[\{U(y) - C\}\{f''(y) - \alpha^2 f(y)\} - U''(y)f(y)] = 0, \tag{1}$$
$$f(1) = f(-1) = f'(1) = f'(-1) = 0.$$

Of the constants involved in this, α has already been defined and R is the Reynold's number. C is a complex parameter for characteristic values of which the boundary problem admits of solution. In terms of such a solution the stream function of the associated flow is given by the formula $f(y)e^{i\alpha(Ct-z)}$. This flow is accordingly stable if and only if the imaginary component of the characteristic value C is positive.

In the special case of a parabolic flow, that is, $U(y)=1-y^2$, the problem has recently been considered by Pekeris.[3] His deductions apply in the main to cases in which the Reynold's number is suitably small, it being shown in particular that stability always obtains if this number does not exceed the smaller of the numbers 2α and α^3. The developments of the smaller characteristic values in series of powers of αR are analyzed, and it is also shown that a symmetrical flow is stable in the limiting case as $\alpha \to 0$. The case of more general velocity profiles $U(y)$ has been considered, among others by Tollmien.[4] His work has special reference to cases in which the Reynold's number is large enough to make $1/\alpha R$ generally negligible, and is focused upon characteristic values with imaginary components that are small.

The present discussion differs widely from those cited. It is of an asymptotic nature and is concerned with the characteristic values for which the product $\alpha R|C|$ is large. The results may be regarded, therefore, as applying to flows associated with large characteristic values if αR is small or moderate, and to characteristic values which are large, moderate, or even small if αR is large. The values of α and R, which are real and positive, are regarded as either fixed or variable. In the latter case the ranges of α, and of $1/\alpha R$, are to be taken as bounded. It will be shown that any flow for which these hypotheses are fulfilled, and for which the imaginary component of C is not very small, is stable.

In the complex plane of the parameter C, let the region consisting of the infinite strip

$$0 < \Re(C) < 1, \qquad \Im(C) > -0, \tag{2}$$

plus an arbitrarily small but fixed neighborhood of the segment of the

[3] Loc. cit. Reference is made to these papers for further citations of the literature.

[4] W. Tollmien, *Ein allgemeines Kriterium der Instabilität laminarer Geschwindigkeitsverteilungen*, Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, 1935, p. 79.

axis of reals from 0 to 1, be shaded, as is shown in Fig. 1. It will be assumed throughout the pending discussion that C lies in the unshaded region, and it will be shown that under this assumption no

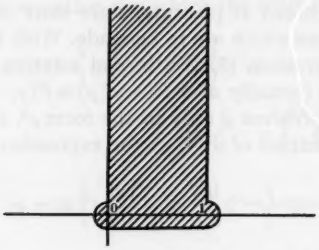

FIG. 1

solution of the boundary problem (1) exists. The inference will then be clear that all characteristic values lie in the shaded region.

In terms of the abbreviations

$$\rho = |\alpha RC|^{1/2} \exp\{i(\pi/4 + (1/2)\arg C)\},$$

$$\phi(y) = [1 - U(y)/C]^{1/2}, \qquad \phi(1) = 1, \tag{3}$$

$$\kappa(y) = U''(y)/C,$$

the differential equation (1) takes the form

$$f^{iv} - \{\rho^2\phi^2(y) + 2\alpha^2\}f'' + \{\rho^2[\alpha^2\phi^2(y) - \kappa(y)] + \alpha^4\}f = 0. \tag{4}$$

The range of the variable is $-1 \leqq y \leqq 1$. Since $U(y)$ is real and takes only values from 0 to 1, whereas C is located in the unshaded portion of the complex plane, the functions $\kappa(y)$, $\phi(y)$, $1/\phi(y)$ are bounded.

If an expression

$$\exp\left\{\rho \int_0^y \phi(y)dy\right\} P(y, \rho), \tag{5}$$

in which $P(y, \rho)$ is an undetermined power series in $1/\rho$, that is,

$$P(y, \rho) \equiv p_0(y) + p_1(y)/\rho + p_2(y)/\rho^2 + \cdots, \tag{6}$$

is substituted for f, the left-hand member of (4) takes the form of an exponential multiplied by a power series in $1/\rho$ with coefficients that are functions of y. By equating these coefficients to zero a sequence of differential equations is obtained which involve the functions $p_j(y)$, $j=0, 1, 2, \cdots$, and from which these functions are successively determinable by integrations. The constants which enter by

virtue of these integrations may be chosen, and are to be chosen, so that the functions $p_j(y)$ with even subscripts are even functions of y, and those with odd subscripts are odd functions of y, and so that $p_0(1)=1$. Those constants which remain undetermined by these stipulations may be chosen at pleasure, since their choice has no effect upon the deductions which are to be made. With the determinations described, the expression (5) is a formal solution of the differential equation (4), and formally also $P(-y, \rho)\equiv P(y, -\rho)$. Inasmuch as the equation (4) involves ρ only in the form ρ^2, it is clear also that with this determination of $P(y, \rho)$ the expression

$$\exp\left\{-\rho\int_0^y \phi(y)dy\right\}P(y, -\rho) \tag{7}$$

is likewise a formal solution of (4).

If, now, in a similar manner, an undetermined series

$$S(y, \rho) \equiv s_0(y) + s_1(y)/\rho^2 + s_2(y)/\rho^4 + \cdots$$

is substituted for f, the left-hand member of (4) takes the form of a power series in $1/\rho^2$. Equating its coefficients to zero leads to the sequence of relations

$$\begin{aligned} &\phi^2 s_0'' - (\alpha^2\phi^2 - \kappa)s_0 = 0, \\ &\phi^2 s_j'' - (\alpha^2\phi^2 - \kappa)s_j = s_{j-1}^{\text{iv}} - 2\alpha^2 s_{j-1}'' + \alpha^4 s_{j-1}, \quad j = 1, 2, 3, \cdots. \end{aligned} \tag{8}$$

These equations may be solved by functions $s_j(y)$ determined successively as odd functions of y, say as functions $s_{j,0}(y)$, $j=0, 1, 2, \cdots$. They may, however, also be solved by functions determined successively as even functions of y, say as functions $s_{j,e}(y)$, $j=0, 1, 2, \cdots$. If this is done, the resulting series

$$\begin{aligned} S_0(y, \rho) &\equiv s_{0,0}(y) + s_{1,0}(y)/\rho^2 + s_{2,0}(y)/\rho^4 + \cdots, \\ S_e(y, \rho) &\equiv s_{0,e}(y) + s_{1,e}(y)/\rho^2 + s_{2,e}(y)/\rho^4 + \cdots \end{aligned} \tag{9}$$

are respectively odd and even, and are formal solutions of the equation (4).

The function $U(y)$ takes on only real values from 0 to 1. For C located anywhere in the unshaded part of the plane in Fig. 1, it accordingly follows that

$$-3\pi/2 \leqq \arg\,[C - U(y)] \leqq \pi/2,$$

and hence, by (3), that

$$-\pi/2 \leqq \arg\,\rho\phi(y) \leqq \pi/2. \tag{10}$$

This, however, is a familiar sufficient condition[5] that the formal expressions (5), (7), and (9), which are in general divergent, have the property of each representing asymptotically some solution of the differential equation (4). There exist, therefore, solutions $f_j(y, \rho)$, $j=1, 2, 3, 4$, of this equation, for which the representations

$$
\begin{aligned}
f_1(y,\rho) &\sim \exp\left\{\rho\int_0^y \phi(y)dy\right\}P(y,\rho) \\
&\quad - \exp\left\{-\rho\int_0^y \phi(y)dy\right\}P(y,-\rho), \\
f_2(y,\rho) &\sim \exp\left\{\rho\int_0^y \phi(y)dy\right\}P(y,\rho) \\
&\quad + \exp\left\{-\rho\int_0^y \phi(y)dy\right\}P(y,-\rho), \\
f_3(y,\rho) &\sim S_0(y,\rho), \qquad f_4(y,\rho) \sim S_e(y,\rho)
\end{aligned}
\tag{11}
$$

are valid. Of these solutions f_1 and f_3 are odd functions of y, and f_2 and f_4 are even.

Any solution of the boundary problem (1) is either even or odd; that is, the flow associated with it is either symmetrical or anti-symmetrical. Let it be assumed, to begin with, that there exists such an *odd* solution. Since it satisfies the differential equation (1), it may be represented in terms of the solutions (11) in the manner $c_1f_1(y)+c_3f_3(y)$, with coefficients c_1, c_3 which are independent of y, and, since it satisfies the boundary conditions (1), this and its derivative vanish at $y=1$. Inasmuch as c_1 and c_3 are not both zero if the solution is non-trivial, it follows that the determinant

$$
\begin{vmatrix} f_1(1,\rho) & f_3(1,\rho) \\ f_1'(1,\rho) & f_3'(1,\rho) \end{vmatrix}
$$

must be zero, a relation which in virtue of the formulas (11), and the evaluations $\phi(1)=p_0(1)=1$, takes the form

$$
\exp\left\{2\rho\int_0^1 \phi(y)dy\right\} \cdot \left\{\frac{[\rho P(1,\rho)+P'(1,\rho)]S_0(1,\rho)-P(1,\rho)S_0'(1,\rho)}{[-\rho P(1,\rho)+P'(1,\rho)]S_0(1,\rho)-P(1,\rho)S_0'(1,\rho)}\right\} \sim 1.
$$

[5] H. L. Turrittin, *Asymptotic solutions of certain ordinary differential equations* · · ·, American Journal of Mathematics, vol. 58 (1936), p. 364.

This, it is found, is reducible to the form

$$\rho\int_0^1 \phi(y)dy + A_1/\rho + A_2/\rho^2 + \cdots \sim B_n i, \tag{12}$$

in which n is an appropriate integer, and in which

$$A_1 = p_0'(1) - s_{0,0}'(1)/s_{0,0}(1), \qquad B_n = (n + 1/2)\pi,$$

if for the case in hand $s_{0,0}(1) \neq 0$, while

$$A_1 = -s_{1,0}(1)/s_{0,0}'(1), \qquad B_n = n\pi,$$

if $s_{0,0}(1) = 0$ (and hence $s_{0,0}'(1) \neq 0$). Since $|\rho|$ is assumed to be large, the numerical values of n and B_n are evidently also large.

Let M be chosen as any sufficiently large real positive constant. The values of C may be considered then in the categories $|C| \geqq M$, and $|C| < M$, and it is convenient to do this. If $|C| \geqq M$, the expansions

$$\phi(y) = 1 - U(y)/2C + \cdots, \qquad A_1 = a_0 + a_1/C + \cdots$$

may be used, and simple calculations show that the constant a_0 is real. The relation obtained by squaring the members of (12) is then found to be

$$C - \int_0^1 U(y)dy - 2ia_0/\alpha R + O(1/C) + O(1/\rho) = B_n^2 i/\alpha R.$$

To satisfy this the value C must evidently have a positive imaginary component, and, if M has been chosen suitably large, a real component which differs so little from the value $\int_0^1 U(y)dy$ that it lies between the values 0 and 1. The relation is, therefore, impossible for the values of C here in question since they lie in the unshaded part of the plane of Fig. 1.

If $|C| < M$ on the other hand, let y_1 and y_2 be chosen in any way such that $0 < y_1 < y_2 < 1$ and $U(y_1) < 1$, $U(y_2) > 0$, and let ϵ designate the smaller of the values $1 - U(y_1)$, $U(y_2)$. Then since C lies in the unshaded part of the plane, it follows that when $y_1 \leqq y \leqq y_2$ then

$$-\pi - \tan^{-1} M/\epsilon \leqq \arg [C - U(y)] < \tan^{-1} M/\epsilon,$$

namely that

$$|\arg \rho\phi(y)| \leqq \pi/2 - (1/2)\cot^{-1} M/\epsilon.$$

Thus, on the interval (y_1, y_2)

$$\Re(\rho\phi(y)) \geqq |\rho\phi(y)| \sin((1/2)\cot^{-1} M/\epsilon),$$

and since by (10) this real part is not negative even outside the interval concerned, it follows that

$$\Re\left\{\rho\int_0^1 \phi(y)dy\right\} \geqq |\rho| \sin\left((1/2)\cot^{-1} M/\epsilon\right)\int_{y_1}^{y_2} |\phi(y)|\, dy.$$

The leading term in the left-hand member of (12) accordingly has a real part which in magnitude is of the order of $|\rho|$. Inasmuch as the right-hand member is pure imaginary, the relation is clearly impossible.

It has thus been shown that all characteristic values for which the flow of the fluid is anti-symmetrical must lie in the shaded part of the plane of Fig. 1. In like manner the same result may be obtained for the characteristic values associated with symmetrical flows. Except for possible flows in the immediate border region between stability and instability, that is, for which the imaginary component of C is very small, it must be inferred, therefore, that any flow for which $\alpha R|C|$ is large is stable.

UNIVERSITY OF WISCONSIN

A NOTE ON HERMITIAN FORMS[1]

N. JACOBSON

In this note we effect a reduction of the theory of hermitian forms of two particular types (coefficients in a quadratic field or in a quaternion algebra with the usual anti-automorphism) to that of quadratic forms. The main theorem (§2) enables us to apply directly the known results on quadratic forms. This is illustrated in the discussion in §3 of a number of special cases.

Let Φ be an arbitrary quasi-field of characteristic different from 2 in which an involutorial anti-automorphism $\alpha \to \bar{\alpha}$ is defined. For the present we do not exclude the cases where Φ is commutative and $\bar{\alpha} \equiv \alpha$ or Φ is a quadratic field with $\alpha \to \bar{\alpha}$ as its automorphism. Suppose $\mathfrak{R}$ is an n-dimensional vector space over Φ. We define a bilinear form (x, y) as a function of pairs of vectors with values in Φ, such that

$$(1)\quad \begin{aligned} (x_1 + x_2, y) &= (x_1, y) + (x_2, y), \quad & (x, y_1 + y_2) &= (x, y_1) + (x, y_2), \\ (x, y\alpha) &= (x, y)\alpha, & (x\alpha, y) &= \bar{\alpha}(x, y), \end{aligned}$$

for all x, y in $\mathfrak{R}$ and α in Φ. If $x_1, x_2, \cdots, x_n$ is a basis for $\mathfrak{R}$ and $(x_i, x_j) = \alpha_{ij}$, the matrix $A = (\alpha_{ij})$ is called the matrix of (x, y) relative to this basis. By (1) it determines (x, y) as $\sum \bar{\xi}_i \alpha_{ij} \eta_j$, if $x = \sum x_i \xi_i$ and $y = \sum x_i \eta_i$. If $y_1, y_2, \cdots, y_n$ where $y_i = \sum x_j \rho_{ji}$ is a second basis for $\mathfrak{R}$ where $R = (\rho_{ij})$ is nonsingular, the matrix of (x, y) relative to this basis is $\overline{R}'AR$. We call A and $\overline{R}'AR$ cogredient. The form (x, y) is hermitian (skew-hermitian), if $(y, x) = \overline{(x, y)}$ ($(y, x) = -\overline{(x, y)}$). This is equivalent to the condition $\overline{A}' = A$ ($\overline{A}' = -A$).

It is readily seen that we may pass from the basis y_i to the x's by a sequence of substitutions of the following two types:

I. $y_i \to y_i$, $(i \neq r)$, $y_r \to y_r + y_s\theta$, $(s \neq r)$.

II. $y_i \to y_i$, $(i \neq r)$, $y_r \to y_r\theta$, $(\theta \neq 0)$.

It follows that we may pass from a matrix to any other matrix cogredient to it by a sequence of transformations of the corresponding types:

I. Addition of the sth column multiplied on the right by θ to the rth together with addition of the sth row multiplied on the left by $\bar{\theta}$ to the rth.

II. Multiplication of the rth column on the right by $\theta \neq 0$ together with multiplication of the rth row on the left by $\bar{\theta}$.

We showed in an earlier paper that any hermitian form or skew-

[1] Presented to the Society, October 28, 1939.

hermitian form with $\bar{\alpha} \neq \alpha$ has a matrix in diagonal form; that is, there is a basis $u_1, u_2, \cdots, u_n$ for $\mathfrak{R}$ such that[2] $(u_i, u_j) = 0$, if $i \neq j$. We call a basis of this type orthogonal and u, v orthogonal, if $(u, v) = 0$. If $(u_i, u_i) = \beta_i \neq 0$ for $i \leqq r$ and $(u_i, u_i) = 0$ for $i > r$, we obtain the diagonal matrix

$$[\beta_1, \beta_2, \cdots, \beta_r, 0, \cdots, 0] \tag{2}$$

for our form.[3] The element β_1 may be taken to be any nonzero element represented by the form, that is, any element for which a u_1 exists such that $(u_1, u_1) = \beta_1$, β_2 is any element represented by (x, y) restricted to the space of vectors orthogonal to u_1, and so on. We note also that β_i may be replaced by $\bar{\gamma}_i \beta_i \gamma_i$, $(\gamma_i \neq 0)$.

The space $\mathfrak{R}_0$ generated by $u_{r+1}, u_{r+2}, \cdots, u_n$ may be characterized as the totality of vectors z, such that $(x, z) = 0$ for all x. The space $\mathfrak{R}_1$ generated by $u_1, u_2, \cdots, u_r$ satisfies the condition $\mathfrak{R} = \mathfrak{R}_0 + \mathfrak{R}_1$, $\mathfrak{R}_0 \cap \mathfrak{R}_1 = 0$. If $\mathfrak{R}_2$ is a second space of this sort, it has a basis of the form $u_i + z_i$, $(i = 1, \cdots, r)$, and hence the matrices of (x, y) in $\mathfrak{R}_1$ and in $\mathfrak{R}_2$ are cogredient. We may therefore restrict our attention to nondegenerate forms $(\mathfrak{R}_0 = 0)$ and shall do so in the remainder of this note.

Two nondegenerate forms $(x, y)_1$ and $(x, y)_2$ in $\mathfrak{R}$ and $\mathfrak{R}'$ respectively are *cogredient* if there is a (1-1) correspondence $x \to x'$ between $\mathfrak{R}$ and $\mathfrak{R}'$ such that[4] $(x, y)_1 = (x', y')_2$. It follows that

$$(x', (y_1 + y_2)')_2 = (x', y_1' + y_2')_2$$

and hence that $(y_1 + y_2)' = y_1' + y_2'$. Similarly $(y\alpha)' = y'\alpha$ and so $x \to x'$ is a linear transformation and $\mathfrak{R}$ and $\mathfrak{R}'$ have the same dimensionality. If $x_1, x_2, \cdots, x_n$ is a basis for $\mathfrak{R}$, then $x_1', x_2', \cdots, x_n'$ is one for $\mathfrak{R}'$. The matrix of $(x, y)_1$ relative to the first basis is the same as that of $(x', y')_2$ relative to the second. Hence the matrices of $(x, y)_1$ and $(x', y')_2$ relative to any bases are cogredient and conversely cogredience of the matrices implies that of the forms.

We shall suppose from now on that Φ is either a quadratic field $\Phi_0(i)$, $i^2 = -\lambda$ and $\bar{\alpha} = \alpha_0 - i\alpha_1$ for $\alpha = \alpha_0 + i\alpha_1$ or that $\Phi = \Phi_0(i, j)$ is a quaternion algebra in which $i^2 = -\lambda$, $j^2 = -\mu$, $k = ij = -ji$ and $\bar{\alpha} = \alpha_0 - i\alpha_1 - j\alpha_2 - k\alpha_3$ for $\alpha = \alpha_0 + i\alpha_1 + j\alpha_2 + k\alpha_3$. We suppose also that (x, y) is hermitian. Then $(x, x) \in \Phi_0$ and any β in (2) may be replaced

[2] *Simple Lie algebras over a field of characteristic zero*, Duke Mathematical Journal, vol. 4 (1938), p. 542.

[3] The above notation for diagonal matrices will be used throughout this note.

[4] $\mathfrak{R}$ and $\mathfrak{R}'$ have the same quasi-field and anti-automorphism.

by $\beta N(\gamma)$, $N(\gamma)=\bar{\gamma}\gamma$. Let Φ_0' be the multiplicative group of nonzero elements in Φ_0, Φ_0^* the subgroup of norms, and $\Gamma=\Phi_0'/\Phi_0^*$. A determinant for any hermitian matrix A has been defined by E. H. Moore.[5] We recall that, if a matrix B has the form (2) with $r=n$, then $\det B=\beta_1\beta_2\cdots\beta_n$ and, if $A=\overline{R}'BR$, $\det A=N(\rho)\det B$. Thus the coset of $\det A$ in Γ is an invariant of the class of matrices cogredient to A (or an invariant of the form). We shall call this coset the discriminant of A (or of the form).

$\mathfrak{R}$ may be regarded as a vector space of $2n$ or $4n$ dimensions over Φ_0 and

$$\{x, y\} = (1/2)[(x, y) + (y, x)] = (1/2)\operatorname{tr}(x, y) \tag{3}$$

is a symmetric form in $\mathfrak{R}$ over Φ_0. The symmetric form $\{x, y\}$ satisfies the special condition

$$\{x\alpha, y\alpha\} = \{x, y\}N(\alpha), \tag{4}$$

whence

$$\{x\bar{\alpha}, y\} = \{x\bar{\alpha}, y\bar{\alpha}^{-1}\bar{\alpha}\} = \{x, y\bar{\alpha}^{-1}\}N(\alpha) = \{x, y\alpha\}.$$

Hence, if $\bar{\alpha}=-\alpha$, $\{x, x\alpha\}=-\{x\alpha, x\}=0$. Conversely, if $\{x, y\}$ is any symmetric bilinear form in $\mathfrak{R}$ over Φ_0 such that (4) holds, (x, y) defined by

$$(x, y) = \begin{cases} \{x, y\} - (i/\lambda)\{x, yi\}, & \text{if } \Phi = \Phi_0(i), \\ \{x, y\} - (i/\lambda)\{x, yi\} - (j/\mu)\{x, yj\} \\ \qquad - (k/\lambda\mu)\{x, yk\}, & \text{if } \Phi = \Phi_0(i, j), \end{cases} \tag{5}$$

is hermitian in $\mathfrak{R}$ over Φ. The relation between (x, y) and $\{x, y\}$ is a reciprocal one and $\{x, y\}$ is nondegenerate if (x, y) is.[6]

Evidently, if $(x, y)_1$ in $\mathfrak{R}$ over Φ and $(x', y')_2$ in $\mathfrak{R}'$ over Φ are cogredient, then $\{x, y\}_1$ and $\{x', y'\}_2$ are cogredient also. Suppose now that $\{x, y\}_1$ and $\{x', y'\}_2$ are cogredient. Then we have u_1 and u_1', such that[7] $(u_1, u_1)_1=\{u_1, u_1\}_1=\{u_1', u_1'\}_2=(u_1', u_1')_2=\beta_1\neq 0$. Let $\mathfrak{R}_1$ and $\mathfrak{R}_1'$ respectively denote the spaces of vectors orthogonal to u_1 and u_1' relative to $(x, y)_1$ and $(x', y')_2$. The space $\mathfrak{R}_1$ may also be characterized as the set of vectors orthogonal to u_1, u_1i if $\Phi=\Phi_0(i)$, or to u_1, u_1i, u_1j, u_1k, if $\Phi=\Phi_0(i, j)$, with respect to $\{x, y\}_1$. A similar

[5] *General Analysis*, I, American Philosophical Society Publication, Philadelphia, 1935.

[6] We make use of the relation $a=\alpha_0+i\alpha_1+j\alpha_2+k\alpha_3=(1/2)[\operatorname{tr} a-(i/\lambda)\operatorname{tr} ai-(j/\mu)\operatorname{tr} aj-(k/\lambda\mu)\operatorname{tr} ak]$.

[7] There exists a vector u_1 such that $(u_1, u_1)\neq 0$. Cf. Jacobson, loc. cit.

characterization holds for $\mathfrak{R}_1'$. The matrix of $\{x, y\}_1$ relative to u_1, $u_1 i$ or $u_1, u_1 i, u_1 j, u_1 k$ and of $\{x', y'\}_2$ relative to u_1', $u_1' i$ or u_1', $u_1' i$, $u_1' j$, $u_1' k$ is

$$[\beta_1, \lambda\beta_1] \quad \text{or} \quad [\beta_1, \lambda\beta_1, \mu\beta_1, \lambda\mu\beta_1]. \tag{6}$$

Hence it follows from a theorem of Witt[8] that $\{x, y\}_1$ and $\{x', y'\}_2$ are cogredient when restricted to $\mathfrak{R}_1$ and $\mathfrak{R}_1'$. By induction $(x, y)_1$ and $(x', y')_2$ are cogredient. Thus we have proved the following theorem:

THEOREM. *A necessary and sufficient condition that two hermitian forms $(x, y)_1$ and $(x, y)_2$ be cogredient is that the corresponding symmetric forms $\{x, y\}_1$ and $\{x, y\}_2$ be cogredient.*

If $u_1, u_2, \cdots, u_n$ is an orthogonal basis, $(u_i, u_i) = \beta_i$, then $u_1, u_1 i, u_2, u_2 i, \cdots, u_n, u_n i$ or $u_1, u_1 i, u_1 j, u_1 k, \cdots, u_n, u_n i, u_n j, u_n k$ is an orthogonal basis for $\mathfrak{R}$ over Φ relative to $\{x, y\}$ and the corresponding matrix, where B_i is as in (6), is

$$[B_1, B_2, \cdots, B_n]. \tag{7}$$

We consider now some special cases:

(1) $\Phi_0(i)$, *where Φ_0 is a field in which every nondegenerate symmetric form in 5 or more variables is a null-form.* Examples of such fields are (a) any p-adic field, (b) an algebraic function field of one variable over a finite constant field.[9] In these cases any nondegenerate symmetric form in 4 or more variables represents every $\alpha \neq 0$ in Φ_0. For, if $\{x, y\}$ represents 0, say $\{u, u\} = 0$, we choose v such that $\{u, v\} = \beta \neq 0$. Then $\{u\xi + v\eta, u\xi + v\eta\} = \eta(2\beta\xi + \gamma\eta)$, $\gamma = \{v, v\}$ and the equation $\eta(2\beta\xi + \gamma\eta) = \alpha$ can be solved for ξ, η in Φ_0. If $\{x, y\}$ does not represent 0, we form the vector space of $(n+1)$ dimensions by adjoining z to $\mathfrak{R}$, and define $\{x\xi + z\eta, x\rho + z\sigma\} = \{x, x\}\xi\rho - \alpha\eta\sigma$. Since this form represents 0, we have $\{x, x\}\xi^2 - \alpha\eta^2 = 0$ for $\eta \neq 0$ since $\{x, x\} \neq 0$. Thus $\{x\xi\eta^{-1}, x\xi\eta^{-1}\} = \alpha$. It follows that any hermitian form in a space of 2 or more dimensions represents any α in Φ_0. Hence we may choose $\beta_1 = \beta_2 = \cdots = \beta_{n-1} = 1$ in (2). Thus two forms are cogredient, if, and only if, they have the same discriminant.

(2) $\Phi_0(i, j)$, Φ_0 *of the same type as in case* (1). Here we may take $\beta_1 = \cdots = \beta_n = 1$ and hence all nondegenerate forms are cogredient.

(3) $\Phi_0(i)$, Φ_0 *a real closed field.* Here we may suppose $\lambda = 1$ and we may suppose $\beta_1 = \cdots = \beta_p = 1$, $\beta_{p+1} = \cdots = \beta_n = -1$. For $\{x, y\}$ we

[8] *Theorie der quadratischen Formen in beliebigen Körpern*, Journal für die reine und angewandte Mathematik, vol. 176 (1936–1937), p. 34.

[9] Witt, loc. cit., p. 40, and Albert, *Quadratic null forms over a function field*, Annals of Mathematics, (2), vol. 39 (1938), pp. 494–505.

obtain $2p$ values $+1$ and $(2n-2p)$ values -1 in the diagonal form. Since the signature is an invariant for bilinear forms it is invariant also for the hermitian form (x, y).

(4) $\Phi_0(i, j)$, Φ_0 *a real closed field.* The considerations are similar to case (3). We find that two nondegenerate hermitian forms are cogredient if and only if they have the same signatures.[10]

(5) $\Phi_0(i)$, Φ_0 *an algebraic number field.* As is well known, the symmetric forms $\{x, y\}_1$ and $\{x, y\}_2$ in $\mathfrak{R}$ over Φ_0 are cogredient, if, and only if, they are cogredient in every p-adic extension of Φ_0. Suppose first that p is a prime spot such that $(-\lambda/p)=1$, that is, $-\lambda$ is a square in the p-adic field[11] Φ_{0p}. Then the matrix B_i in (7) is cogredient in Φ_{0p} to $[\beta_i, -\beta_i]$ and hence also[12] to $[1, -1]$. Thus any two matrices of the form (7) are cogredient. If $(-\lambda/p)=-1$, $\Phi_{0p}(i)$ is a quadratic field over Φ_{0p}. Hence $\{x, y\}_1$ and $\{x, y\}_2$ are cogredient, if, and only if, $(x, y)_1$ and $(x, y)_2$ are cogredient in $\mathfrak{R}$ over $\Phi_{0p}(i)$. The condition for this is that the discriminants be the same when p is finite and the signatures be the same when p is infinite. Combining these results, we see that a necessary and sufficient condition that two nondegenerate hermitian forms in $\mathfrak{R}$ over Φ be cogredient is that they have the same discriminant and the same signature at the infinite prime spots for which λ is positive.[13]

(6) $\Phi_0(i, j)$, Φ_0 *an algebraic number field.* To obtain conditions for cogredience of $(x, y)_1$ and $(x, y)_2$ we again consider $\{x, y\}_1$ and $\{x, y\}_2$. Let p be a prime spot at which $((-\lambda, -\mu)/p)=1$, that is, $\Phi_{0p}(i, j)$ is a matrix algebra. Then either $-\lambda$ is a square in Φ_{0p} or $-\mu$ is a norm in $\Phi_{0p}(i)$. In the first case B_i is cogredient to $[\beta_i, -\beta_i, \mu\beta_i, -\mu\beta_i]$ and hence to $[1, -1, 1, -1]$. If $-\mu$ is a norm in $\Phi_{0p}(i)$, the bilinear form with matrix $[\beta_i, \lambda\beta_i, \mu\beta_i]$ represents 0 and hence is cogredient to $[1, -1, -\lambda\mu\beta_i]$, and again (6) is cogredient to (8). If p is a prime spot for which $\Phi_{0p}(i, j)$ is a division algebra, $(x, y)_1$ and $(x, y)_2$ are always cogredient, if p is finite, and these forms are cogredient for p infinite, if, and only if, they have the same signatures. Thus a necessary and sufficient condition that $(x, y)_1$ and $(x, y)_2$ in $\mathfrak{R}$ over Φ be cogredient is that these forms have the same signatures for all infinite prime spots for which $\Phi_{0p}(i, j)$ is a division algebra.

UNIVERSITY OF NORTH CAROLINA

[10] E. H. Moore, loc. cit., p. 193.

[11] See Witt and the references cited there to Hasse's papers.

[12] This is a consequence of Witt's theorem that any two symmetric forms in two variables which are nonsingular and represent 0 are cogredient (Witt, p. 34).

[13] Cf. Landherr, *Äquivalenz Hermitescher Formen über einem beliebigen algebraischen Zahlkörper*, Abhandlungen aus dem mathematischen Seminar der Hansischen Universität, vol. 11 (1936), pp. 245–248.

AN EXTENSION OF A COVARIANT DIFFERENTIATION PROCESS[1]

MARIE M. JOHNSON

Craig[2] has considered tensors $T^{\alpha\cdots}_{\beta\cdots}$ whose components are functions of n variables represented by x and their m derivatives $x', x'', \cdots, x^{(m)}$. He obtained the covariant derivative

$$T^{\alpha\cdots}_{\beta\cdots x^{(m-1)}\gamma} - mT^{\alpha\cdots}_{\beta\cdots x^{(m)}\lambda}\left\{{\lambda \atop \gamma}\right\}, \qquad m \geqq 2, \tag{1}$$

where

$$\left\{{\lambda \atop \gamma}\right\} \equiv x'^{\alpha}\Gamma^{\lambda}_{\gamma\alpha} + (1/2)x''^{\beta}f_{\gamma\lambda\beta}f^{\beta\lambda}, \tag{2}$$

and partial differentiation in (1) is denoted by the added subscript. Throughout, a repeated letter in one term indicates a sum of n terms. The purpose of this note is to derive another tensor from $T^{\alpha\cdots}_{\beta\cdots}$ whose covariant rank is one larger. The general process will be shown clearly by using $T^{\alpha}(x, x', x'', x''')$.

The extended point transformation

$$x^{\alpha} = x^{\alpha}(y), \qquad x'^{\alpha} = \frac{\partial x^{\alpha}}{\partial y^{i}}\, y'^{i},$$

$$x''^{\alpha} = \frac{\partial x^{\alpha}}{\partial y^{i}}\, y''^{i} + \frac{\partial^2 x^{\alpha}}{\partial y^{i}\partial y^{j}}\, y'^{i}y'^{j}, \cdots, \qquad \alpha = 1, \cdots, n,$$

gives the tensor equations of transformation of the tensor T^{α} as

$$\overline{T}^{i}(y, y', y'', y''') = T^{\alpha}(x, x', x'', x''')\partial y^{i}/\partial x^{\alpha}, \tag{3}$$

where y stands for the n variables $y^1, y^2, \cdots, y^n$ and a similar notation is used for the derivatives y', y'', and y'''. On differentiating equations (3) with respect to y'^{k} it is found that

$$\overline{T}^{i}_{y'^{k}} = \left(T^{\alpha}_{x'\beta}\frac{\partial x'^{\beta}}{\partial y'^{k}} + T^{\alpha}_{x''\beta}\frac{\partial x''^{\beta}}{\partial y'^{k}} + T^{\alpha}_{x'''\beta}\frac{\partial x'''^{\beta}}{\partial y'^{k}}\right)\partial y^{i}/\partial x^{\alpha}. \tag{4}$$

The derivatives can be expressed by using the following general formulas:

---

[1] Presented to the Society, April 15, 1939.

[2] H. V. Craig, *On a covariant differentiation process*, this Bulletin, vol. 37 (1931), pp. 731–734.

$$(5)\qquad \frac{\partial x^{(m-1)\beta}}{\partial y^{(m-2)k}} = (m-1)\frac{\partial x'^{\beta}}{\partial y^{k}}, \qquad \frac{\partial x^{(m)\beta}}{\partial y^{(m-2)k}} = \frac{m(m-1)}{2}\frac{\partial x''^{\beta}}{\partial y^{k}},$$

in which $\partial x'^{\beta}/\partial y^{k}$ are eliminated by[3]

$$(6)\qquad \overline{\left\{ {i \atop k} \right\}}\frac{\partial x^{\beta}}{\partial y^{i}} = \frac{\partial x'^{\beta}}{\partial y^{k}} + \left\{ {\beta \atop \gamma} \right\}\frac{\partial x^{\gamma}}{\partial y^{k}}.$$

The derivatives $\partial x''^{\beta}/\partial y^{k}$ are simplified by first writing

$$(7)\qquad x''^{\beta} = \frac{\partial x^{\beta}}{\partial y^{j}}y''^{j} + \overline{\Gamma}^{r}_{jk}y'^{j}y'^{k}\frac{\partial x^{\beta}}{\partial y^{r}} - \Gamma^{\beta}_{\alpha\delta}x'^{\alpha}x'^{\delta},$$

with the help of (2), (6) and[4] $f_{\alpha\beta\gamma}x'^{\beta}=0$. It is necessary also to have

$$(8)\qquad \frac{\partial^{2}x^{\beta}}{\partial y^{j}\partial y^{k}} = \overline{\Lambda}^{i}_{jk}\frac{\partial x^{\beta}}{\partial y^{i}} - \Lambda^{\beta}_{\alpha\delta}\frac{\partial x^{\alpha}}{\partial y^{j}}\frac{\partial x^{\delta}}{\partial y^{k}},$$

where

$$\Lambda^{\beta}_{\alpha\delta} = \Gamma^{\beta}_{\alpha\delta} - (1/2)f^{\beta\gamma}(f_{\delta\gamma\tau}\left\{ {\tau \atop \alpha} \right\} + f_{\gamma\alpha\tau}\left\{ {\tau \atop \delta} \right\} - f_{\alpha\delta\tau}\left\{ {\tau \atop \gamma} \right\}).$$

This is obtained from Taylor's[5] formula (19) in the following way. Multiply this formula by $(\partial y^{k}/\partial x^{\epsilon})f^{\beta\epsilon}=(\partial x^{\beta}/\partial y^{l})\bar{f}^{kl}$, and sum for k. Use the tensor equations for $f_{\alpha\beta\gamma}$ and substitute from (6) for $\partial x'^{\gamma}/\partial y^{j}$.

By means of formulas (6) and (8) and the tensor $Q^{\beta}(x, x', x'') \equiv x''^{\beta}+\Gamma^{\beta}_{\alpha\delta}x'^{\alpha}x'^{\delta}$ the partial derivatives of (7) have the form

$$(9)\qquad \frac{\partial x''^{\beta}}{\partial y^{k}} = -\left| {\beta \atop \gamma} \right|\frac{\partial x^{\gamma}}{\partial y^{k}} + \overline{\left| {r \atop k} \right|}\frac{\partial x^{\beta}}{\partial y^{r}} - 2\left\{ {\beta \atop \alpha} \right\}\overline{\left\{ {i \atop k} \right\}}\frac{\partial x^{\alpha}}{\partial y^{i}} + 2\overline{\left\{ {r \atop i} \right\}}\overline{\left\{ {i \atop k} \right\}}\frac{\partial x^{\beta}}{\partial y^{r}},$$

in which we have the nontensor form

$$(10)\qquad \left| {\beta \atop \gamma} \right| = Q^{\beta}_{x^{\gamma}} - Q^{\beta}_{x'^{\alpha}}\left\{ {\alpha \atop \gamma} \right\} + Q^{\alpha}\Lambda^{\beta}_{\alpha\gamma}.$$

If formulas (6) and (9) are substituted in equations (5) and the results used in (4), we find

$$\overline{T}^{i}_{y'^{k}} = (T^{\alpha}_{x'^{\beta}} - 2T^{\alpha}_{x'^{\delta}}\left\{ {\delta \atop \beta} \right\} - 3T^{\alpha}_{x'^{\delta}}\left| {\delta \atop \beta} \right|)\frac{\partial x^{\beta}}{\partial y^{k}}\frac{\partial y^{i}}{\partial x^{\alpha}}$$
$$-(-2\overline{T}^{i}_{y'^{l}}\overline{\left\{ {l \atop k} \right\}} - 3\overline{T}^{i}_{y'^{l}}\overline{\left| {l \atop k} \right|}).$$

[3] J. H. Taylor, *A generalization of Levi-Civita's parallelism and the Frenet formulas*' Transactions of this Society, vol. 27 (1925), p. 255.

[4] J. H. Taylor, loc. cit., p. 248.

[5] J. H. Taylor, loc. cit., p. 254.

Hence the new tensor whose covariant rank has been increased by one is

$$T^{\alpha}_{x,\beta} - 2T^{\alpha}_{x',\delta}\left\{{\delta \atop \beta}\right\} - 3T^{\alpha}_{x'',\delta}\left|{\delta \atop \beta}\right|, \tag{11}$$

where $\left\{{\delta \atop \beta}\right\}$ and $\left|{\delta \atop \beta}\right|$ are defined in (2) and (10).

Because of the general relations in (5) it is easy to verify that the tensor

$$T^{\alpha\cdots}_{\gamma\cdots x^{(m-2)}\beta} - (m-1)T^{\alpha\cdots}_{\gamma\cdots x^{(m-1)}\delta}\left\{{\delta \atop \beta}\right\} - \frac{m(m-1)}{2}T^{\alpha\cdots}_{\gamma\cdots x^{(m)}\delta}\left|{\delta \atop \beta}\right|, \tag{12}$$
$$m \geqq 3,$$

has a covariant rank which is one larger than that of $T^{\alpha\cdots}_{\gamma\cdots}$ whose components are functions of $(x, x', \cdots, x^{(m)})$.

If the components of the tensor $T^{\alpha}(x, x', x'', x''')$ do not contain the derivatives x''', then (11) reduces to Craig's covariant derivative (1), and if there are no x'' or x''' derivatives, then the result is a partial differentiation with respect to x'.

The usual rules for the derivative of a sum of tensors of the same type and rank and for the product of any tensors are preserved by this provess.

If $m=2$, a scalar $T(x, x', x'')$ will give a covariant tensor which is similar to that in (11) when the tensor equations for $\overline{T}(y, y', y'')$ are differentiated with respect to y instead of y'. The tensor is

$$T_{,\beta} - T_{x,\delta}\left\{{\delta \atop \beta}\right\} - T_{x',\delta}\left|{\delta \atop \beta}\right|. \tag{13}$$

However, if $m=2$ and a tensor $T^{\alpha}(x, x', x'')$ is used, an extra term $T^{\delta}\Lambda^{\alpha}_{\delta\beta}$ has to be added to three terms similar to those in (13). If this process is performed on the tensor $Q^{\alpha}(x, x', x'')$, the result is the zero tensor.

OBERLIN COLLEGE

INDEFINITELY DIFFERENTIABLE FUNCTIONS WITH PRESCRIBED LEAST UPPER BOUNDS[1]

IRWIN E. PERLIN

1. **Introduction.** Let $F(x)$ be a real indefinitely differentiable function of the real variable x defined on the interval $a \leqq x \leqq b$, and let M_n denote the least upper bound of $|F^{(n)}(x)|$ on that interval.[2] In this paper we shall establish sufficient conditions that there exist an indefinitely differentiable function $F(x)$ taking on certain prescribed M_n.

It is easy to see that M_0 and M_1 can be assigned arbitrarily. However, the first three upper bounds M_0, M_1, and M_2 must satisfy certain inequalities.[3] Let us consider the interval $(0, 1)$. Let t_1 be the value of x for which $|F^{(1)}(x)|$ attains its maximum. Then

$$F(1) - F(t_1) = (1 - t_1)F^{(1)}(t_1) + (1/2!)F^{(2)}(\theta_1)(1 - t_1)^2,$$

where $t_1 < \theta_1 < 1$. And similarly

$$F(0) - F(t_1) = -t_1F^{(1)}(t_1) + (1/2!)F^{(2)}(\theta_2)t_1^2,$$

where $0 < \theta_2 < t_1$. On subtracting these equations we obtain

$$F^{(1)}(t_1) = F(1) - F(0) + (1/2!)\{F^{(2)}(\theta_2)t_1^2 - F^{(2)}(\theta_1)(1 - t_1)^2\},$$

$$M_1 \leqq 2M_0 + M_2/2!.$$

By the same procedure we can obtain for the interval $(0, a)$

$$M_1 \leqq 2M_0/a + M_2a/2!. \tag{1}$$

In the case of the interval $(0, \infty)$ we can replace (1) by a more precise inequality. Since a is arbitrary, we can replace a by the positive value which minimizes the right side of (1), and obtain

$$M_1 \leqq 2(M_0M_2)^{1/2}.$$

Ore[4] in a recent paper employed the results of W. Markoff to obtain certain inequalities connecting the least upper bounds of $|F^{(i)}(x)|$, $(1 \leqq i \leqq n)$, with those of $|F(x)|$ and $|F^{(n+1)}(x)|$ where $F(x)$ is a function with bounded $(n+1)$th derivative. For the first derivative the inequality (1) is slightly better than that obtained by Ore.

---

[1] Presented to the Society, September 8, 1939.

[2] By $F^{(0)}(x)$ we shall mean $F(x)$.

[3] Hadamard, Comptes Rendus des Séances de la Société Mathématique de France, 1914, pp. 68–72; T. Carlman, *Les fonctions quasi analytiques*, Paris, 1926.

[4] O. Ore, *On functions with bounded derivatives*, Transactions of this Society, vol. 43 (1938), pp. 321–326.

2. Construction of an indefinitely differentiable function with prescribed least upper bounds. We now prove the following theorem.

THEOREM. *If the sequence* $\{M_i a^i\}$ *is monotone decreasing, then there exists an indefinitely differentiable function* $F(x)$ *defined on* $(0 \leqq x \leqq a)$ *such that* M_i *is the least upper bound of* $|F^{(i)}(x)|$ *on* $(0 \leqq x \leqq a)$.

Define

$$0 \leqq S_i = \sum_{j=0}^{\infty} (-1)^j \frac{M_{i+j} a^j}{j!} \leqq M_i, \qquad F(x) = \sum_{i=0}^{\infty} S_i \frac{x^i}{i!}.$$

Now the function $F(x)$ so defined is an entire function. Let $x = b$. Then

$$F(b) = \sum_{i=0}^{\infty} S_i \frac{b^i}{i!} = \sum_{i=0}^{\infty} (-1)^i \frac{M_i}{i!} \left(\sum_{j=0}^{i} (-1)^j C_{i-j,j} a^{i-j} b^j \right),$$

$$= \sum_{i=0}^{\infty} (-1)^i \frac{M_i (a-b)^i}{i!}.$$

And since the series

$$\sum_{i=0}^{\infty} \frac{M_i a^i \left| (1 - b/a)^i \right|}{i!} < M_0 \sum_{i=0}^{\infty} \frac{\left| (1 - b/a)^i \right|}{i!}$$

converges, the series $\sum_{i=0}^{\infty} (S_i/i!) b^i$ converges.[5] Further,

$$F^{(i)}(a) = \sum_{j=0}^{\infty} (-1)^j M_{i+j} a^j \left(\sum_{k=0}^{j} \frac{(-1)^k}{k!(j-k)!} \right) = M_i,$$

and

$$\left| F^{(i)}(x) \right| \leqq F^{(i)}(a) = M_i, \qquad 0 \leqq x \leqq a.$$

Thus we have given explicitly a function $F(x)$ satisfying the conditions of the theorem.

ARMOUR INSTITUTE OF TECHNOLOGY

[5] See T. J. I'A. Bromwich, *Theory of Infinite Series*, London, 1931, p. 266.

A DECOMPOSITION OF ADDITIVE SET FUNCTIONS[1]

R. S. PHILLIPS

This paper is concerned with a decomposition theorem for additive functions on an additive family of sets to either real numbers or a Banach space. Additive bounded set functions have as yet been little studied. However the recent paper of Hildebrandt[2] illustrates their importance.

We shall use the following notation:

(a) T: an abstract class of arbitrary elements t.

(b) $\mathfrak{Z}$: a completely additive family of subsets τ of T; that is, $T \in \mathfrak{Z}$, $\tau \in \mathfrak{Z}$ implies $T-\tau \in \mathfrak{Z}$, and $\tau_n \in \mathfrak{Z}$ for $n=1, 2, \cdots$ implies $\sum\tau_n \in \mathfrak{Z}$.

(c) α: a set function on $\mathfrak{Z}$ to real numbers.

(d) A: the subclass of set functions on $\mathfrak{Z}$ to real numbers which are additive and bounded; that is, $\tau_1, \tau_2 \in \mathfrak{Z}$ and $\tau_1 \cdot \tau_2 = 0$ implies $\alpha(\tau_1+\tau_2)=\alpha(\tau_1)+\alpha(\tau_2)$.

(e) C: the subclass of set functions on $\mathfrak{Z}$ to real numbers which are completely additive (c.a.), that is, $\tau_n \in \mathfrak{Z}$ for $n=1, 2, \cdots$ and $\tau_i \cdot \tau_j = 0$ if $i \neq j$ implies $\alpha(\sum\tau_n)=\sum\alpha(\tau_n)$. The functions in C are bounded.[3]

The notations A_P and C_P refer to the subclasses of A and C respectively whose elements are nonnegative.

(f) x: a set function on $\mathfrak{Z}$ to a Banach space[4] X. The definitions of additive and c.a. set functions are formally retained. If $\{\tau_n\}$ is a sequence of disjoint sets of $\mathfrak{Z}$ and $x(\tau)$ is c.a., then $\sum x(\tau_n)$ is unconditionally convergent.[5]

(g) C_X: the class of c.a. set functions on $\mathfrak{Z}$ to X.

In the statement of the following theorems, D will designate any one of the classes A, A_P, C, C_P, and $\bar{\tau}$ will denote the cardinal number of τ.

THEOREM 1. *Let $\aleph$ be an infinite cardinal number not greater than $\bar{\bar{T}}$. For every $\alpha \in D$ there exists an unique decomposition $\alpha=\alpha_1+\alpha_2$ and a set $R(\alpha) \in \mathfrak{Z}$ of cardinal number not greater than $\aleph$ such that $\alpha_1, \alpha_2 \in D$,*

[1] Presented to the Society April 15, 1939, under the title *On additive set functions.*

[2] T. H. Hildebrandt, *On bounded linear functional operations*, Transactions of this Society, vol. 36 (1934), pp. 868–875.

[3] S. Saks, *Theory of the Integral*, Monografje Matematyczne, Warsaw, 1937, p. 10, Theorem 6.1.

[4] S. Banach, *Théorie des Opérations Linéaires*, Monografje Matematyczne, Warsaw, 1932, chap. 5.

[5] If x_n is a series of elements of X and if every subseries $\sum x_n$ is convergent, then $\sum x_n$ is said to be unconditionally convergent.

$$\alpha_1(\tau) = \alpha(R\cdot\tau), \qquad \alpha_2(\tau) = 0 \quad \text{if} \quad \bar{\tau} \leqq \aleph .$$

Let $\Sigma \equiv E_\tau[\tau \in \mathfrak{Z}, \bar{\tau} \leqq \aleph, \alpha(\tau) \neq 0]$. We define a transfinite sequence $(\tau_1, \tau_2, \cdots; \tau_\omega, \cdots, \tau_\lambda, \cdots)$ as follows: τ_1 is an arbitrary element of Σ. Suppose τ_λ have been defined for all $\lambda < \mu$. If there exists τ such that $\tau \cdot \sum_{\lambda<\mu}\tau_\lambda = 0$ and $\tau \in \Sigma$, then we set $\tau = \tau_\mu$.

As $\alpha(\tau)$ is bounded, $\alpha(\tau)$ cannot differ from zero on a nondenumerable number of disjoint sets. The sequence therefore contains only a denumerable set of elements.

Let $R = \sum_\lambda \tau_\lambda$. Then $R \in \mathfrak{Z}$ and $\overline{R} \leqq \aleph$. We define $\alpha_1(\tau) = \alpha(R\cdot\tau)$, $\alpha_2(\tau) = \alpha(\tau) - \alpha_1(\tau) = \alpha(\tau - R\cdot\tau)$. The $\alpha_1(\tau)$, $\alpha_2(\tau)$ are clearly elements of D. If $\bar{\tau} \leqq \aleph$, then by the definition of R, $\alpha_2(\tau) = \alpha(\tau - R\cdot\tau) = 0$.

Although the set R is not unique, the function decomposition is unique: Suppose there exist two different sets R_1, R_2 having the properties of the R defined above. The set identity $R_1\cdot\tau + (R_2 - R_1)\cdot\tau = R_2\cdot\tau + (R_1 - R_2)\cdot\tau$ and $\alpha[(R_1 - R_2)\cdot\tau] = 0 = \alpha[(R_2 - R_1)\cdot\tau]$ imply that $\alpha(R_1\cdot\tau) = \alpha(R_2\cdot\tau)$.

A set function α on $\mathfrak{Z}$ will be said to be *nonsingular* if for every $t \in \mathfrak{Z}$, $\alpha(t) = 0$. A set function α on $\mathfrak{Z}$ will be called $\aleph$*-homogeneous* if there exists a set R such that $R \in \mathfrak{Z}$, $\overline{R} = \aleph$, $\alpha(\tau) = \alpha(R\cdot\tau)$, and $\alpha(\tau) = 0$ if $\bar{\tau} < \aleph$.

Without loss of generality we may consider only nonsingular set functions because for every $\alpha \in D$ there exists a unique decomposition $\alpha = \alpha_1 + \alpha_2$ and a denumerable set $\{t_i\}$ of elements of T, such that $\alpha_1, \alpha_2 \in D$, $\alpha_1(\tau) = \sum_{i=1}^{\infty}\alpha(\tau\cdot t_i)$, and α_2 is nonsingular. We omit the proof.

THEOREM 2. *For every nonsingular $\alpha \in D$, there exists an unique decomposition $\alpha = \sum_i \alpha_i$, the sum being absolutely convergent, and such that α_i is $\aleph_i$-homogeneous and $\aleph_i \neq \aleph_j$ if $i \neq j$.*

In the proof of this theorem an induction is made on the infinite cardinals not exceeding that of T, well-ordered according to magnitude. We define a transfinite sequence of set functions $(\alpha_1, \alpha_2, \cdots; \alpha_\omega, \cdots, \alpha_\lambda, \cdots)$ as follows: Suppose α_λ have been defined for all $\lambda < \mu$ and (1) only a denumerable number of the α_λ are not identically zero; (2) $\sum_{\lambda \leqq \lambda_0}|\alpha_\lambda(\tau)| < \infty$; and (3) $\alpha_\lambda \in D$ and is $\aleph_\lambda$-homogeneous. By Theorem 1 there exist $R_\mu \in \mathfrak{Z}$ and a decomposition $\alpha = \alpha_\mu^1 + \alpha_\mu^2$ such that $\overline{R}_\mu \leqq \aleph_\mu$, $\alpha_\mu^1(\tau) = \alpha(R_\mu\cdot\tau)$, $\alpha_\mu^2(\tau) = 0$ if $\bar{\tau} \leqq \aleph_\mu$, and $\alpha_\mu^1, \alpha_\mu^2 \in D$. Clearly $\alpha_\lambda(\tau) = \alpha(R_\mu\cdot R_\lambda\cdot\tau)$ if $\lambda < \mu$.

Let $\alpha_\mu(\tau) = \alpha_\mu^1(\tau) - \sum_{\lambda<\mu}\alpha_\lambda(\tau)$. We consider the following cases:

I. $\alpha \in C$, C_P. Let $\overline{R}_\mu = R_\mu - \sum_{\tau_\mu}\overline{R}_\lambda$ where $\pi_\mu \equiv E_\lambda[\lambda < \mu, \alpha_\lambda \not\equiv 0]$. The sets $\overline{R}_\mu$ are disjoint. Suppose $\alpha_\lambda(\tau) = \alpha(\overline{R}_\lambda\cdot\tau)$ for $\lambda < \mu$. Then by (1)

$$\alpha_\mu(\tau) = \alpha(R_\mu \cdot \tau) - \sum_{\tau_\mu} \alpha_\lambda(\tau) = \alpha(R_\mu \cdot \tau) - \sum_{\tau_\mu} \alpha(R_\mu \cdot \overline{R}_\lambda \cdot \tau)$$

$$= \alpha\left[\left(R_\mu - \sum_{\tau_\mu} R_\mu \cdot \overline{R}_\lambda\right) \cdot \tau\right] = \alpha(\overline{R}_\mu \cdot \tau).$$

It is clear that (1), (2), and (3) are satisfied for $\mu+1$. $\alpha_\lambda \not\equiv 0$ implies that $\alpha(\tau) \not\equiv 0$ for some subset of $\overline{R}_\lambda$. As the $\overline{R}_\lambda$ are disjoint, the sequence will contain only a denumerable number of functions not identically zero.

II. $\alpha \in A_P$. For $\lambda_0 < \mu$, $\alpha(T) \geqq a^1_{\lambda_0}(T) \equiv \sum_{\lambda \leqq \lambda_0} \alpha_\lambda(T) \geqq \sum_{\lambda \leqq \lambda_0} \alpha_\lambda(\tau)$. Clearly (1) and (2) are satisfied for $\mu+1$, and the sequence contains only a denumerable number of functions not identically zero. Let λ_i be a spanning sequence for $E_\lambda[\lambda < \mu, \alpha_\lambda \not\equiv 0]$. Then

$$\alpha_\mu(\tau) = \alpha^1_\mu(\tau) - \sum_{\lambda < \mu} \alpha_\lambda(\tau) = \alpha(R_\mu \cdot \tau) - \lim_{i \to \infty} \alpha^1_{\lambda_i}(\tau)$$

$$= \alpha(R_\mu \cdot \tau) - \lim_{i \to \infty} \alpha(R_\mu \cdot R_{\lambda_i} \cdot \tau).$$

Hence (3) is likewise satisfied.

III. $\alpha \in A$. Every $\alpha \in A$ has a decomposition $\alpha = \alpha_1 - \alpha_2$ where $\alpha_1, \alpha_2 \in A_P$. An application of II to α_1 and α_2 gives the desired decomposition.

The decomposition is unique: Any two sequences of homogeneous functions differ in a first function, α_μ. But this is contrary to $\alpha^1_\mu = \sum_{\lambda \leqq \mu} \alpha_\lambda$ being unique.

In these theorems the restriction that the additive bounded set function be defined over an additive family $\mathfrak{Z}$ is optional, since the range of definition of such a function can always be extended to an additive family. The type of argument used by Pettis[6] will prove this statement.

We next consider the possibility of extending these theorems to functions $x(\tau)$ on $\mathfrak{Z}$ to a Banach space. The theorem is not in general valid for additive bounded set functions of this type. This is illustrated by $x(\tau)$ defined on all subsets of $T \equiv (0, 1)$ to the space X of bounded functions on $S \equiv (0, 1)$ where $x(\tau)$ is the characteristic function of the subset of S which has the same coordinate values as τ. Clearly there exists no denumerable set R such that $x(\tau - R\tau) = 0$ for all denumerable sets τ.

However analogous theorems are obtained for c.a. set functions on $\mathfrak{Z}$ to X.

[6] B. J. Pettis, *Linear functionals and completely additive set functions*, Duke Mathematical Journal, vol. 4 (1938), p. 554, Theorem 1.1.

THEOREM 3. *Let* $\aleph$ *be an infinite cardinal number not greater than* $\overline{\overline{T}}$. *For every* $x \in C_X$ *there exists a unique decomposition* $x = x_1 + x_2$ *and a set* $R(x) \in \mathfrak{Z}$ *of cardinal power not greater than* $\aleph$ *such that* $x_1, x_2 \in C_X$, $x_1(\tau) = x(R \cdot \tau)$, $x_2(\tau) = 0$ *if* $\overline{\tau} \leqq \aleph$.

$x(\tau) \neq 0$ on at most a denumerable number of disjoint sets of $\mathfrak{Z}$. Suppose the contrary. Then there exists a denumerable sequence of disjoint sets $\{\tau_i\}$ and an $e > 0$ such that $\|x(\tau_i)\| > e$, $(i = 1, 2, \cdots)$. As $x(\tau)$ is c.a., $\sum_i x(\tau_i)$ converges. The supposition is therefore false.

The argument used in Theorem 1 will now prove the theorem.

THEOREM 4. *For every nonsingular* $x \in C_X$, *there exists an unique decomposition* $x = \sum_i x_i$, *the sum being unconditionally convergent, and such that* x_i *is* $\aleph_i$*-homogeneous and* $\aleph_i \neq \aleph_j$ *if* $i \neq j$.

The proof is identical with that of I in Theorem 2. Again there will exist disjoint $\overline{R}_\mu$'s such that $x_\mu(\tau) = x(\overline{R}_\mu \cdot \tau)$.

UNIVERSITY OF MICHIGAN

ON THE COMPLETENESS OF A CERTAIN METRIC SPACE WITH AN APPLICATION TO BLASCHKE'S SELECTION THEOREM[1]

G. BALEY PRICE

1. **Introduction.** The purpose of this note is to prove that the metric space whose elements are the closed, bounded, non-null subsets of a complete metric space, and whose metric is the Hausdorff distance, is complete; and, using this result and others already known, to give a simple proof of Blaschke's selection theorem.

2. **Preliminaries.** Let K be a metric space with elements $x, y, \cdots$ and distance function $d(x, y)$. A sequence $x_1, x_2, \cdots$ in K such that $\sum_1^\infty d(x_i, x_{i+1})$ converges has been called an absolutely convergent sequence by MacNeille[2] [7, p. 192]. Every absolutely convergent sequence is a Cauchy sequence, and every Cauchy sequence contains absolutely convergent subsequences.

Let K^* be a metric space whose elements $X, Y, \cdots$ are the closed, bounded, and non-null subsets of K, and whose distance function $D(X, Y)$ is the Hausdorff distance between the sets X and Y (see Hausdorff [5, pp. 145–146] and Kuratowski [6, pp. 89–90]).

3. **The theorem.** *If K is complete, then K^* is also complete.*

Let $X_1, X_2, \cdots$ be any Cauchy sequence in K^*; without loss of generality we can assume that it is absolutely convergent. We shall define a set X and show that it is the limit of the given sequence. Let x_1 be any point in X_1, x_2 any point in X_2 such that $d(x_1, x_2) < D(X_1, X_2) + 2^{-1}$, x_3 any point in X_3 such that $d(x_2, x_3) < D(X_2, X_3) + 2^{-2}$, and so on. The existence of points $x_2, x_3, \cdots$ with the properties stated follows from the definition of the Hausdorff distance. Every point x_i in X_i is a member of a sequence $x_1, x_2, \cdots$ of the kind described. The sequence $x_1, x_2, \cdots$ is absolutely convergent and hence a Cauchy sequence; since K is complete, it has a limit x_0 in K. Let X_0 be the locus of all the points x_0 obtained as the limits of all possible sequences formed in the manner stated; let X be the closure of X_0. Then X is closed, bounded, and non-null, and X is in K^*. We shall show that $\lim X_k = X$. Let any $\epsilon > 0$ be given. Choose $n = n(\epsilon)$ so that $\sum_n^\infty [D(X_i, X_{i+1}) + 2^{-i}] < \epsilon/2$. Let $x^* \in X$, and let x_0 be the limit of a

[1] Presented to the Society, December 28, 1938, under the title *Spaces whose elements are sets*.

[2] Numbers in square brackets refer to the references at the end.

sequence $x_1, x_2, \cdots, x_k, \cdots$ and such that $d(x^*, x_0) < \epsilon/2$. Then the distance from x^* to X_k is equal to or less than

$$d(x^*, x_0) + \sum_k^\infty d(x_i, x_{i+1}) < d(x^*, x_0) + \sum_k^\infty [D(X_i, X_{i+1}) + 2^{-i}]$$
$$< \epsilon/2 + \epsilon/2 = \epsilon$$

if $k \geqq n$. Since every point x_k in X_k belongs to a sequence $x_1, x_2, \cdots$, the distance from any point x_k to X, which does not exceed the distance from x_k to the limit x_0 of the sequence $x_1, x_2, \cdots, x_k, \cdots$, is equal to or less than

$$\sum_k^\infty d(x_i, x_{i+1}) < \sum_k^\infty [D(X_i, X_{i+1}) + 2^{-i}] < \epsilon/2$$

if $k \geqq n$. From these facts it follows that $D(X_k, X) < \epsilon$ for $k \geqq n$, and hence that $\lim X_k = X$. Thus the (absolutely convergent) Cauchy sequence $X_1, X_2, \cdots$ in K^* has the limit X in K^*, and the proof of the theorem is complete.

4. **The space K^* when K is a Banach space.** The space K^* has additional properties when K is a Banach space, that is, a space which is linear, normed, and complete (see Banach [1, p. 53]). Let aX denote the set of elements ax, $x \in X$, when a is a real number; let $X+Y$ denote the set of elements $x+y$, $x \in X$ and $y \in Y$; let $C[X]$ denote the closed convex extension of X; and let $\rho(X)$ denote the diameter of X. Then K^* has, in addition to its elementary properties as a metric space, the following ones:

$$D(aX, aY) = |a| D(X, Y) \quad \text{for every real number } a; \tag{4.1}$$

$$D(X_1 + \cdots + X_n, Y_1 + \cdots + Y_n) \leqq D(X_1, Y_1) + \cdots + D(X_n, Y_n); \tag{4.2}$$

$$D(C[X], C[Y]) \leqq D(X, Y); \tag{4.3}$$

$$D(X + Y_1, X + Y_2) \leqq D(Y_1, Y_2); \tag{4.4}$$

$$\rho(X_i) \leqq \rho(X_1 + X_2) \leqq \rho(X_1) + \rho(X_2), \quad i = 1, 2; \tag{4.5}$$

$$\rho(C[X]) = \rho(X), \quad D(C[X], 0) = D(X, 0). \tag{4.6}$$

The last two of these relations have been given by Birkhoff [2, pp. 368, 360]. The proofs of the others will be given elsewhere. It can be shown by means of examples that the inequality may hold in (4.4).

If the limit of a sequence $X_1, X_2, \cdots$ of convex sets in K^* is a set X, it follows from (4.3) that X also is convex. For

$D(X_i, C[X]) = D(C[X_i], C[X]) \leqq D(X_i, X)$; and since $D(X_i, X) \to 0$, $\lim X_i = C[X]$. But since a sequence has a unique limit, we have $C[X] = X$, and X is convex.

5. **Blaschke's selection theorem.** Let E be a closed and compact subset of a Banach space K, and let E^* denote the subset of K^* which consists of the closed, non-null subsets of E. Then both E and E^* are totally bounded, and E^* is closed and compact in K^* (see Hausdorff [5, pp. 107–108] and Kuratowski [6, p. 91]). Let E_C^* denote the subset of E^* which consists of convex sets. Since E^* is totally bounded, any infinite set of elements in $E^* \subset E^*$ contains a Cauchy sequence; since K^* is complete and E^* is closed, this sequence has a limit in E^*. By the result at the end of the last section, this limit element is itself a closed, convex set and therefore belongs to E_C^*. We have thus shown that E^* is closed and compact. This result is Blaschke's selection theorem extended to a Banach space (see Blaschke [3] and Bonnesen and Fenchel [4, p. 34]).

References

1. S. Banach, *Théorie des Opérations Linéaires*, Monografje Matematyczne, vol. 1, Warsaw, 1932.

2. G. Birkhoff, *Integration of functions with values in a Banach space*, Transactions of this Society, vol. 38 (1935), pp. 357–378.

3. W. Blaschke, *Kreis und Kugel*, Leipzig, Veit, 1916.

4. T. Bonnesen and W. Fenchel, *Theorie der konvexen Körper*, Ergebnisse der Mathematik und ihrer Grenzgebiete, vol. 3, no. 1, 1934.

5. F. Hausdorff, *Mengenlehre*, Berlin and Leipzig, de Gruyter, 2d edition, 1927.

6. C. Kuratowski, *Topologie* I, Monografje Matematyczne, vol. 3, Warsaw, 1933.

7. H. M. MacNeille, *Extensions of measure*, Proceedings of the National Academy of Sciences, vol. 24 (1938), pp. 188–193.

University of Kansas

SOME PROBLEMS IN INTERPOLATION BY CHARACTERISTIC FUNCTIONS OF LINEAR DIFFERENTIAL SYSTEMS OF THE FOURTH ORDER

KATHARINE E. O'BRIEN

In this paper we consider the convergence to $f(x)$, defined on $[0, 1]$, of

$$\Sigma_p[f(x)] = \alpha_{0p}u_0(x) + \alpha_{1p}u_1(x) + \cdots + \alpha_{pp}u_p(x),$$

where $u_n(x)$, $(n=0, 1, \cdots, p)$, are characteristic functions of certain self-adjoint linear differential systems of fourth order,

$$\alpha_{np} = \sum_{k=0}^{p}{}' f(x_k)u_n(x_k)\left\{\sum_{k=0}^{p}{}' u_n^2(x_k)\right\}^{-1}, \qquad n = 0, 1, \cdots, p,$$

and the symbol $\sum'$ is used in the sense $\sum_{k=0}^{p}{}' y_k = y_0/2 + \sum_{k=1}^{p} y_k$. Throughout the discussion, $x_k = 2k/(2p+1)$, $(k=0, 1, \cdots, p)$. The differential systems considered are

$$u^{(iv)} - \rho^4 u = 0,$$

with boundary conditions

I. $u'(0)=0$, $u'''(0)=0$, $u'(1)=0$, $u'''(1)+u(1)=0$,
II. $u'(0)=0$, $u'''(0)=0$, $u'(1)+u(1)=0$, $u'''(1)+u''(1)=0$,
III. $u(0)=0$, $u''(0)=0$, $u(1)=0$, $u''(1)+u'(1)=0$,
IV. $u'(0)=0$, $u'''(0)=0$, $u(1)=0$, $u''(1)+u'(1)=0$,
V. $u(0)=0$, $u'(0)=0$, $u(1)=0$, $u'(1)=0$,
VI. $u'(0)=0$, $u'''(0)=0$, $u(1)=0$, $u'(1)=0$.

The following theorems may be proved for these systems respectively.

I, II. *If $f(x)$ is continuous and of bounded variation in $[0, 1]$, then $\lim_{p\to\infty}\Sigma_p[f(x)]=f(x)$ uniformly in $[0, 1]$.*

III. *If $f(x)$ is continuous and of bounded variation in $[0, 1]$ and $f(0)=f(1)=0$, then $\lim_{p\to\infty}\Sigma_p[f(x)]=f(x)$ uniformly in $[0, 1]$.*

IV. *If $f(x)$ is continuous and of bounded variation in $[0, 1]$ and $f(1)=0$, then $\lim_{p\to\infty}\Sigma_p[f(x)]=f(x)$ uniformly in $[0, 1-\eta]$.*

V, VI. *If $f(x)$ satisfies a Lipschitz condition in $[0, 1]$ and $f(0)=f(1)=0$, then $\lim_{p\to\infty}\Sigma_p[f(x)]=f(x)$ uniformly in $[\eta, 1-\eta]$.*

Here and hereafter $\eta>0$ is arbitrarily small but fixed.

The method of proof for these theorems, as well as for those to fol-

low, is essentially the same as that of C. M. Jensen [1], who considered convergence properties of a somewhat similar sum using Sturm-Liouville functions.

The following equivalence theorems (the term "equivalence" is used in the sense of Jensen) may be proved for systems I, II, III, IV, respectively.

I, II. *If $f(x)$ is continuous in $[0, 1]$, then*

$$\lim_{p\to\infty} \{\Sigma_p[f(x)] - T_p[f(x)]\} = 0$$

uniformly in $[0, 1]$, where $T_p[f(x)]$ is the cosine interpolation formula

$$T_p[f(x)] = a_{0p} + a_{1p}\cos \pi x + \cdots + a_{pp}\cos p\pi x,$$

$$a_{np} = \sum_{k=0}^{p}{}' f(x_k)\cos n\pi x_k \left\{ \sum_{k=0}^{p}{}' \cos^2 n\pi x_k \right\}^{-1}$$

$$= \begin{cases} \dfrac{4}{2p+1}\displaystyle\sum_{k=0}^{p}{}' f(x_k)\cos n\pi x_k, & n = 1, 2, \cdots, p, \\ \dfrac{2}{2p+1}\displaystyle\sum_{k=0}^{p}{}' f(x_k), & n = 0. \end{cases}$$

III. *If $f(x)$ is continuous in $[0, 1]$, then*

$$\lim_{p\to\infty} \{\Sigma_p[f(x)] - \overline{T}_p[f(x)]\} = 0$$

uniformly in $[\eta, 1-\eta]$, where $\overline{T}_p[f(x)]$ is the sine interpolation formula

$$\overline{T}_p[f(x)] = \bar{a}_{1p}\sin \pi x + \bar{a}_{2p}\sin 2\pi x + \cdots + \bar{a}_{pp}\sin p\pi x,$$

$$\bar{a}_{np} = \sum_{k=0}^{p}{}' f(x_k)\sin n\pi x_k \left\{ \sum_{k=0}^{p}{}' \sin^2 n\pi x_k \right\}^{-1}$$

$$= \frac{4}{2p+1}\sum_{k=0}^{p}{}' f(x_k)\sin n\pi x_k, \qquad n = 1, 2, \cdots, p.$$

IV. *If $f(x)$ is continuous in $[0, 1]$, then*

$$\lim_{p\to\infty} \{\Sigma_p[f(x)] - U_p[f(x)]\} = 0$$

uniformly in $[\eta, 1-\eta]$, where $U_p[f(x)]$ is given by

$$U_p[f(x)] = b_{0p}\cos (\pi/2)x + b_{1p}\cos (3\pi/2)x + \cdots + b_{pp}\cos (p + 1/2)\pi x,$$

$$b_{np} = \sum_{k=0}^{p}{}' f(x_k)\cos (n + 1/2)\pi x_k \left\{ \sum_{k=0}^{p}{}' \cos^2 (n + 1/2)\pi x_k \right\}^{-1}$$

$$= \begin{cases} \dfrac{4}{2p+1}\sum_{k=0}^{p}{}' f(x_k)\cos(n+1/2)\pi x_k, & n = 0, 1, \cdots, p-1, \\ \dfrac{2}{2p+1}\sum_{k=0}^{p}{}' f(x_k)\cos(n+1/2)\pi x_k, & n = p. \end{cases}$$

The first set of theorems also holds using characteristic functions of $u^{(iv)} - [\rho^4+\lambda(x)]u=0$, with the above boundary conditions, $\lambda(x)$ being an arbitrary continuous function. Here, provided ${}_0u(x)\not\equiv 0$, the term ${}_0u(x)\sum_{k=0}^{'p} f(x_k)\,{}_0u(x_k)\{\sum_{k=0}^{'p}{}_0u^2(x_k)\}^{-1}$ is adjoined to $\Sigma_p[f(x)]$, where ${}_0u(x)$ is the characteristic function corresponding to $\rho=0$. (For $\lambda(x)\equiv 0$, $\rho=0$ is not a characteristic number.)

We give here the proof of the theorem for system V. System V is the problem of lateral vibrations of an elastic homogeneous rod clamped at both ends [2]. We first state two lemmas.

LEMMA 1. *If $f(x)$ is continuous and of bounded variation in $[0, 1]$ and $f(0)=f(1)=0$, then $W_p[f(x)]$ and $w_p[f(x)]$ tend to the same limit $M(x)$ uniformly in $[\eta, 1-\eta]$ as $p\to\infty$, where*

$$W_p[f(x)] = c_{0p}[\cos(\pi/2)x - \sin(\pi/2)x] + c_{1p}[\cos(3\pi/2)x - \sin(3\pi/2)x] + \cdots + c_{pp}[\cos(p+1/2)\pi x - \sin(p+1/2)\pi x],$$

$$c_{np} = \frac{2}{2p+1}\sum_{k=0}^{p}{}' f(x_k)[\cos(n+1/2)\pi x_k - \sin(n+1/2)\pi x_k], \qquad n = 0, 1, \cdots, p,$$

$$w_p[f(x)] = c_0[\cos(\pi/2)x - \sin(\pi/2)x] + c_1[\cos(3\pi/2)x - \sin(3\pi/2)x] + \cdots + c_p[\cos(p+1/2)\pi x - \sin(p+1/2)\pi x],$$

$$c_n = \int_0^1 f(t)[\cos(n+1/2)\pi t - \sin(n+1/2)\pi t]dt, \qquad n = 0, 1, \cdots, p.$$

LEMMA 2. *If $f(x)$ satisfies a Lipschitz condition in $[0, 1]$ and $f(0)=f(1)=0$, then there exists a constant C, depending only on η, such that for n sufficiently great, $p\geqq n$, and x in $[\eta, 1-\eta]$,*

$$|\alpha_{np}u_n(x) - c_{np}[\cos(n+1/2)\pi x - \sin(n+1/2)\pi x]| < C/n^2.$$

To prove Lemma 1, write

$$W_p[f(x)] = {}_1W_p[f(x)] + {}_2W_p[f(x)] - {}_3W_p[f(x)],$$
$$w_p[f(x)] = {}_1w_p[f(x)] + {}_2w_p[f(x)] - {}_3w_p[f(x)],$$

where

$$ {}_1W_p[f(x)] = \frac{2}{2p+1}\sum_{k=0}^{p}{}'f(x_k)\sum_{n=0}^{p}\cos(n+1/2)\pi x_k\cos(n+1/2)\pi x, $$

$$ {}_2W_p[f(x)] = \frac{2}{2p+1}\sum_{k=0}^{p}{}'f(x_k)\sum_{n=0}^{p}\sin(n+1/2)\pi x_k\sin(n+1/2)\pi x, $$

$$ {}_3W_p[f(x)] = \frac{2}{2p+1}\sum_{k=0}^{p}{}'f(x_k)\sum_{n=0}^{p}\sin(n+1/2)\pi(x_k+x), $$

$$ {}_1w_p[f(x)] = \int_0^1 f(t)\sum_{n=0}^{p}\cos(n+1/2)\pi t\cos(n+1/2)\pi x dt, $$

$$ {}_2w_p[f(x)] = \int_0^1 f(t)\sum_{n=0}^{p}\sin(n+1/2)\pi t\sin(n+1/2)\pi x dt, $$

$$ {}_3w_p[f(x)] = \int_0^1 f(t)\sum_{n=0}^{p}\sin(n+1/2)\pi(t+x)dt. $$

We first show that if $f(x)$ is continuous and of bounded variation in $[0, 1]$, then $\lim_{p\to\infty} {}_1W_p[f(x)] = (1/2)f(x)$ uniformly in $[0, 1-\eta]$. We employ the cosine interpolation formula $T_p[f(x)]$. We shall use $r_n(x)$ as generic notation for a function uniformly bounded in n and for x in $[0, 1]$, unless the range for x is otherwise stated, and r_n $[r_{np}]$ for a quantity depending on n [n and p] and uniformly bounded in n [n and p]. We have

$$ \begin{aligned} T_p[f(x)] &= \frac{2}{2p+1}\sum_{k=0}^{p}{}'f(x_k) + \frac{4}{2p+1}\sum_{k=0}^{p}{}'f(x_k)\sum_{n=1}^{p}\cos n\pi x_k\cos n\pi x \\ &= \frac{1}{2p+1}\sum_{k=0}^{p}{}'f(x_k)\left[\frac{\sin(p+1/2)\pi(x_k-x)}{\sin(\pi/2)(x_k-x)}\right. \\ &\qquad \left. + \frac{\sin(p+1/2)\pi(x_k+x)}{\sin(\pi/2)(x_k+x)}\right] \\ &= \frac{1}{2p+1}\sum_{k=0}^{p}{}'f(x_k)\left[\frac{\sin p\pi(x_k-x)\cos(\pi/2)(x_k-x)}{\sin(\pi/2)(x_k-x)}\right. \\ &\qquad \left. + \frac{\sin p\pi(x_k+x)\cos(\pi/2)(x_k+x)}{\sin(\pi/2)(x_k+x)}\right] \\ &\quad + \frac{2}{2p+1}\cos p\pi x\sum_{k=0}^{p}{}'f(x_k)\cos p\pi x_k. \end{aligned} $$

Using auxiliary Lemma A stated below, we have

$$T_p[f(x)] = \frac{1}{2p+1}\sum_{k=0}^{p}{}'f(x_k)\left[\frac{\sin p\pi(x_k-x)\cos(\pi/2)(x_k-x)}{\sin(\pi/2)(x_k-x)} + \frac{\sin p\pi(x_k+x)\cos(\pi/2)(x_k+x)}{\sin(\pi/2)(x_k+x)}\right] + \frac{r_p(x)}{p},$$

and

$$\begin{aligned}{}_1W_p[f(x)] &= \frac{2}{2p+1}\sum_{k=0}^{p}{}'f(x_k)\sum_{n=0}^{p-1}\cos(n+1/2)\pi x_k\cos(n+1/2)\pi x \\ &\quad + r_p(x)/p \\ &= \frac{1}{2(2p+1)}\sum_{k=0}^{p}{}'f(x_k)\left[\frac{\sin p\pi(x_k-x)}{\sin(\pi/2)(x_k-x)} + \frac{\sin p\pi(x_k+x)}{\sin(\pi/2)(x_k+x)}\right] + \frac{r_p(x)}{p}.\end{aligned}$$

Thus

$$\begin{aligned}&{}_1W_p[f(x)] - (1/2)T_p[f(x)] \\ &= \frac{1}{2(2p+1)}\sum_{k=0}^{p}{}'f(x_k)\left[\frac{1-\cos(\pi/2)(x_k-x)}{\sin(\pi/2)(x_k-x)}\sin p\pi(x_k-x) + \frac{1-\cos(\pi/2)(x_k+x)}{\sin(\pi/2)(x_k+x)}\sin p\pi(x_k+x)\right] + \frac{r_p(x)}{p} \\ &= \frac{1}{2(2p+1)}\sum_{k=0}^{p}{}'f(x_k)[\tan(\pi/4)(x_k-x)\sin p\pi(x_k-x) \\ &\qquad + \tan(\pi/4)(x_k+x)\sin p\pi(x_k+x)] + r_p(x)/p.\end{aligned}$$

By reason of the nature of Lemma A and the fact that for t in $[0, 1]$ the functions $\tan(\pi/4)(t-x)$ and $\tan(\pi/4)(t+x)$ are of uniform bounded variation with respect to x in $[0, 1-\eta]$, we have ${}_1W_p[f(x)]-(1/2)T_p[f(x)]=r_p(x)/p$ for x in $[0, 1-\eta]$. Hence [3] $\lim_{p\to\infty}{}_1W_p[f(x)]=(1/2)f(x)$ uniformly in $[0, 1-\eta]$.

Employing $t_p[f(x)]$, the partial sum of order p in the Fourier cosine series,

$$t_p[f(x)] = a_0 + a_1\cos\pi x + \cdots + a_p\cos p\pi x,$$

$$a_n = \frac{\int_0^1 f(x)\cos n\pi x dx}{\int_0^1 \cos^2 n\pi x dx} = \begin{cases} 2\int_0^1 f(x)\cos n\pi x dx, & n = 1, 2, \cdots, p, \\ \int_0^1 f(x)dx, & n = 0, \end{cases}$$

we may show similarly that if $f(x)$ is continuous and of bounded variation in $[0, 1]$, then $\lim_{p\to\infty} {}_1w_p[f(x)] = (1/2)f(x)$ uniformly in $[0, 1-\eta]$. Employing $\overline{T}_p[f(x)]$, the sine interpolation formula, and $\bar{t}_p[f(x)]$, the partial sum of order p in the Fourier sine series,

$$\bar{t}_p[f(x)] = \bar{a}_1 \sin \pi x + \bar{a}_2 \sin 2\pi x + \cdots + \bar{a}_p \sin p\pi x,$$

$$\bar{a}_n = \frac{\int_0^1 f(x) \sin n\pi x dx}{\int_0^1 \sin^2 n\pi x dx} = 2\int_0^1 f(x) \sin n\pi x dx,\ n = 1, 2, \cdots, p,$$

we may likewise show that if $f(x)$ is continuous and of bounded variation in $[0, 1]$ and $f(0) = f(1) = 0$, then

$$\lim_{p\to\infty} {}_2W_p[f(x)] = (1/2)f(x), \qquad \lim_{p\to\infty} {}_2w_p[f(x)] = (1/2)f(x)$$

uniformly in $[0, 1-\eta]$. In connection with ${}_1w_p[f(x)]$ and ${}_2w_p[f(x)]$ we use Lemma B, stated below.

Finally, we show that if $f(x)$ is of bounded variation in $[0, 1]$, then ${}_3W_p[f(x)]$ and ${}_3w_p[f(x)]$ both tend to $(1/2)\int_0^1 f(t)\{\sin(\pi/2)(t+x)\}^{-1}dt$ uniformly in $[\eta, 1-\eta]$. We have

$${}_3w_p[f(x)] = \frac{1}{2}\int_0^1 f(t) \frac{1 - \cos(p+1)\pi(t+x)}{\sin(\pi/2)(t+x)} dt.$$

For x in $[\eta, 1-\eta]$,

$${}_3w_p[f(x)] = \frac{1}{2}\int_0^1 \frac{f(t)}{\sin(\pi/2)(t+x)} dt - \frac{1}{2}\int_0^1 f(t) \frac{\cos(p+1)\pi(t+x)}{\sin(\pi/2)(t+x)} dt.$$

Effecting some trigonometric reductions on the second integrand and using Lemma B, we have

$${}_3w_p[f(x)] = \frac{1}{2}\int_0^1 \frac{f(t)}{\sin(\pi/2)(t+x)} dt + \frac{r_p(x)}{p}.$$

We may then prove that

$$\lim_{p\to\infty} {}_3W_p[f(x)] = \frac{1}{2}\int_0^1 \frac{f(t)}{\sin(\pi/2)(t+x)} dt$$

uniformly in $[\eta, 1-\eta]$, using the fact just established that

$$\lim_{p\to\infty} {}_3w_p[f(x)] = \frac{1}{2}\int_0^1 \frac{f(t)}{\sin(\pi/2)(t+x)}\,dt$$

uniformly in $[\eta, 1-\eta]$, together with Lemma C, stated below.

LEMMA A. *If $f(x)$ is of bounded variation in $[0, 1]$, then for $n=1, 2, \cdots, 2p$,*

$$\frac{1}{2p+1}\left|\sum_{k=0}^{p}{}' f(x_k)\cos n\pi x_k\right| = \frac{r_{np}}{n},$$

$$\frac{1}{2p+1}\left|\sum_{k=0}^{p}{}' f(x_k)\sin n\pi x_k\right| = \frac{r_{np}}{n}.$$

Also

$$\frac{1}{2p+1}\left|\sum_{k=0}^{p}{}' f(x_k)\sin(2p+1)\pi x_k\right| = 0.$$

LEMMA B. *If $f(x)$ is of bounded variation in $[0, 1]$, then for $n>0$,*

$$\left|\int_0^1 f(x)\cos n\pi x dx\right| = \frac{r_n}{n}, \qquad \left|\int_0^1 f(x)\sin n\pi x dx\right| = \frac{r_n}{n}.$$

LEMMA C. *If $f(x)$ is of bounded variation in $[0, 1]$, then for any preassigned $\epsilon>0$ there exists Q such that for $p>q\geqq Q$ and x in $[\eta, 1]$,*

$$\frac{2}{2p+1}\left|\sum_{k=0}^{p}{}' f(x_k)\sum_{n=q+1}^{p}\sin(n+1/2)\pi(x_k+x)\right| < \epsilon.$$

Lemma 2 is proved by means of auxiliary Lemmas D, E, and A.

LEMMA D. *For n sufficiently great, and $p\geqq n$,*

$$\left[\sum_{k=0}^{p}{}' u_n^2(x_k)\right]^{-1} = \frac{2}{2p+1}\left(1+\frac{r_{np}}{n}\right).$$

LEMMA E. *If $f(x)$ satisfies a Lipschitz condition in $[0, 1]$ and if $f(0)=f(1)=0$, then for $n=1, 2, \cdots, p$,*

$$\frac{1}{2p+1}\left|\sum_{k=0}^{p}{}' f(x_k)\exp\{-(n+1/2)\pi x_k\}\right| = \frac{r_{np}}{n^2},$$

$$\frac{1}{2p+1}\left|\sum_{k=0}^{p}{}' f(x_k)\exp\{-(n+1/2)\pi(1-x_k)\}\right| = \frac{r_{np}}{n^2}.$$

The following asymptotic expression is known [4] for the characteristic functions of system V:

$$u_n(x) = \cos(n+1/2)\pi x - \sin(n+1/2)\pi x + (-1)^n \exp\{-(n+1/2)\pi(1-x)\} - \exp\{-(n+1/2)\pi x\} + \exp\{-(n+1/2)\pi\} r_n(x).$$

Define

$$\sigma_p[f(x)] = \alpha_0 u_0(x) + \alpha_1 u_1(x) + \cdots + \alpha_p u_p(x),$$

$$\alpha_n = \frac{\int_0^1 f(x)u_n(x)dx}{\int_0^1 u_n^2(x)dx}, \qquad n = 0, 1, \cdots, p.$$

Denote by $\Sigma_p^{(r,s)}$ the sum of the terms in $\Sigma_p[f(x)]$ with subscripts r to s inclusive; similarly in the other sums.

For x in $[\eta, 1-\eta]$,

$$|f(x) - \Sigma_p[f(x)]| \leqq |\sigma^{(0,N)} - \Sigma_p^{(0,N)}| + |W_p^{(N+1,p)} - \Sigma_p^{(N+1,p)}| + |W_p^{(0,N)} - w^{(0,N)}| + |M(x) - W_p^{(0,p)}| + |w^{(0,N)} - M(x)| + |f(x) - \sigma^{(0,N)}|.$$

The right-hand member can be made arbitrarily small by choosing p sufficiently large. Call the six terms $D_1, D_2, \cdots, D_6$. Given ϵ, first choose N sufficiently large so that $D_2 < \epsilon$, $D_5 < \epsilon$, $D_6 < \epsilon$, for $p \geqq N+1$. Having fixed N, choose P sufficiently large so that $D_1 < \epsilon, D_3 < \epsilon, D_4 < \epsilon$ for $p \geqq P$. It remains to justify these statements. For D_6 we use a result in a paper by J. D. Tamarkin [5]. For D_2 we use Lemma 2, and for D_4 and D_5 Lemma 1. In D_1 and D_3 we are dealing essentially with integrals of continuous functions and the sums which tend to the integrals as limits, the number of terms being finite.

Now consider $u^{(iv)} - [\rho^4 + \lambda(x)]u = 0$, $\lambda(x)$ being an arbitrary continuous function. A fundamental system of solutions of $u^{(iv)} - \rho^4 u = 0$ is $u_1 = \cos \rho x$, $u_2 = \sin \rho x$, $u_3 = e^{\rho x}$, $u_4 = e^{-\rho x}$. By the method of variation of constants we have, as an equation satisfied by u,

$$u = A\cos\rho x + B\sin\rho x + Ce^{\rho x} + De^{-\rho x} + \frac{1}{2\rho^3}\int_0^x \lambda(t)u(t)\left[-\sin\rho(x-t) + \frac{e^{\rho(x-t)}}{2} - \frac{e^{-\rho(x-t)}}{2}\right]dt,$$

where A, B, C, D are arbitrary constants. Applying the boundary conditions of system V and choosing the multiplicative constant so that $A=1$, we have, assuming $\rho\neq 0$,

$$u = \cos\rho x - \sin\rho x$$
$$+\Big[\sin\rho x\Big\{-2e^{-\rho}\sin\rho+\frac{1}{2\rho^3}\int_0^1\lambda(t)u(t)\big[-e^{-\rho}\sin\rho(1-t)$$
$$-e^{-\rho}\cos\rho(1-t)+e^{-\rho t}\big]dt\Big\}$$
$$+e^{-\rho(1-x)}\sin\rho-\frac{1}{4\rho^3}\int_0^1\lambda(t)u(t)\big[-e^{-\rho(1-x)}\sin\rho(1-t)$$
$$-e^{-\rho(1-x)}\cos\rho(1-t)+e^{\rho(x-t)}\big]dt$$
$$-e^{-\rho x}+e^{-\rho(1+x)}\cos\rho+\frac{1}{4\rho^3}\int_0^1\lambda(t)u(t)\big[-e^{-\rho(1+x)}\sin\rho(1-t)$$
$$-e^{-\rho(1+x)}\cos\rho(1-t)+e^{-\rho(x+t)}\big]dt\Big]\big\{1-e^{-\rho}\sin\rho-e^{-\rho}\cos\rho\big\}^{-1}$$
$$+\frac{1}{2\rho^3}\int_0^x\lambda(t)u(t)\Big[-\sin\rho(x-t)+\frac{e^{\rho(x-t)}}{2}-\frac{e^{-\rho(x-t)}}{2}\Big]dt,$$

and, as an equation satisfied by the characteristic numbers $\rho=\rho_n$,

$$\cos\rho = 2e^{-\rho}-e^{-2\rho}\cos\rho$$
$$+\frac{\sin\rho}{2\rho^3}\int_0^1\lambda(t)u(t)\Big[e^{-\rho}\cos\rho(1-t)-\frac{e^{-\rho t}}{2}-\frac{e^{-\rho(2-t)}}{2}\Big]dt$$
$$+\frac{\cos\rho}{2\rho^3}\int_0^1\lambda(t)u(t)\Big[-e^{-\rho}\sin\rho(1-t)+\frac{e^{-\rho t}}{2}-\frac{e^{-\rho(2-t)}}{2}\Big]dt$$
$$+\frac{1}{4\rho^3}\int_0^1\lambda(t)u(t)\big[\sin\rho(1-t)-\cos\rho\,(1-t)+e^{-\rho(1-t)}\big]dt$$
$$-\frac{e^{-\rho}}{4\rho^3}\int_0^1\lambda(t)u(t)\big[-e^{-\rho}\sin\rho(1-t)-e^{-\rho}\cos\rho(1-t)+e^{-\rho t}\big]dt.$$

For each $\rho=\rho_n$, u is a continuous function; we show u uniformly bounded in n. In u, the terms

$$-\frac{1}{4\rho^3}\int_0^1\lambda(t)u(t)e^{\rho(x-t)}dt\big\{1-e^{-\rho}\sin\rho-e^{-\rho}\cos\rho\big\}^{-1},$$
$$\frac{1}{4\rho^3}\int_0^x\lambda(t)u(t)e^{\rho(x-t)}dt$$

may be combined to give

$$\left[-\frac{1}{4\rho^3}\int_x^1 \lambda(t)u(t)e^{\rho(x-t)}dt - \frac{\sin\rho+\cos\rho}{4\rho^3}\int_0^x \lambda(t)u(t)e^{-\rho(1-x+t)}dt\right] \cdot\left\{1-e^{-\rho}\sin\rho-e^{-\rho}\cos\rho\right\}^{-1}.$$

For t in $[x, 1]$, $x-t\leqq 0$; for t in $[0, x]$, $1-x+t\geqq 0$. Also, there exists a constant $c>0$ such that $1-e^{-\rho}\sin\rho-e^{-\rho}\cos\rho\geqq c$ for $\rho=\rho_n$. Call $M_n=\max_{[0,1]}|u_n(x)|$, $K=\int_0^1|\lambda(t)|\,dt$. Then

$$|u_n(x)|\leqq 2+\frac{5}{c}+\frac{17M_nK}{4c\rho_n^3}, \qquad M_n\leqq 2+\frac{5}{c}+\frac{17M_nK}{4c\rho_n^3},$$

$$M_n\leqq\frac{2+5/c}{1-17K/(4c\rho_n^3)}.$$

Thus for n sufficiently great, $M_n\leqq 4+10/c$. The remaining n's form a finite set. Hence $u_n(x)$ is uniformly bounded in n. Thus

$$u_n(x)=\cos\rho_nx+\sin\rho_nx+e^{-\rho_n(1-x)}\sin\rho_n-e^{-\rho_nx}+r_n(x)/\rho_n^3,$$

and $\cos\rho_n=\phi(\rho_n)$, where $\lim_{n\to\infty}\phi(\rho_n)=0$. Thus $\rho_n=(n+1/2)\pi+\epsilon_n$, where $\lim_{n\to\infty}\epsilon_n=0$. It follows that

$$\begin{aligned}u_n(x)=&\cos(n+1/2)\pi x-\sin(n+1/2)\pi x\\&+(-1)^n\exp\left\{-(n+1/2)\pi(1-x)\right\}\\&-\exp\left\{-(n+1/2)\pi x\right\}+r_n(x)/n^3.\end{aligned}$$

The theorem for system V follows.

References

1. C. M. Jensen, *Some problems in the theory of interpolation by Sturm-Liouville functions*, Transactions of this Society, vol. 29 (1927), pp. 54–79.

2. Rayleigh, *Theory of Sound*, vol. 1, London, Macmillan, 1929, pp. 275–276.

3. Dunham Jackson, *Theory of Approximation*, American Mathematical Society Colloquium Publications, vol. 11, New York, 1930, p. 134.

4. Stekloff and Tamarkin, *Problème des vibrations transversales d'une verge élastique homogène*, Rendiconti del Circolo Matematico di Palermo, vol. 31 (1911), pp. 341–362.

5. J. D. Tamarkin, *Some general problems of the theory of ordinary linear differential equations and expansion of an arbitrary function in series of fundamental functions*, Mathematische Zeitschrift, vol. 27 (1927), pp. 1–54.

Brown University

AN ALMOST UNIVERSAL FORM

GORDON PALL

P. R. Halmos[1] obtained the 88 possible forms (a, b, c, d), $0<a\leqq b\leqq c\leqq d$, which represent all positive integers with one exception, and proved that property for all except for the form $h=(1, 2, 7, 13)$. A proof for h follows.

The forms $f=(1, 2, 7)$ and $g=(1, 1, 14)$ constitute the reduced forms of a genus.[2] Between them they represent all positive integers not of the form[3] $\Lambda=7^{2k+1}(7m+3, 5, 6)$. The identities

$$x^2+y^2+14z^2=x^2+2((y+7z)/3)^2+7((y-2z)/3)^2$$
$$=y^2+2((x+7z)/3)^2+7((x-2z)/3)^2$$

show that every number represented by g with either $y\equiv -z$ or $x\equiv -z \pmod 3$ is also represented by f. Hence every number $3n$ and $3n+1$ not of the form Λ is represented by f. For, $x\equiv y\equiv 0$, $z\not\equiv 0$, and $x, y\not\equiv 0$, $z\equiv 0 \pmod 3$ both imply $g\equiv 2$. If $N=3n$ or $3n+1$ is of the form Λ, then $7\mid N$, so that $N-13\cdot 3^2\neq\Lambda$. Similarly, one of $3n+2-13$ and $3n+2-52$ is not of the form Λ; but neither of these is congruent to 2 (mod 3). These linear forms are positive if $n\geqq 39$; h represents all integers not less than 119. The only number less than 119 not represented in (1, 2, 7, 13) is found to be 5.

McGILL UNIVERSITY

---

[1] This Bulletin, vol. 44 (1938), pp. 141–144.

[2] See any table of positive ternaries.

[3] For example, see B. W. Jones, Transactions of this Society, vol. 33 (1931), pp. 111–124.

THE GENERAL LIMIT THEOREM IN THE THEORY OF PROBABILITY[1]

WOLFGANG STERNBERG

Let $x_1, x_2, \cdots, x_n$ be independent chance variables with any distribution functions $V_1, V_2, \cdots, V_n$.

We put[2] for all $i=1, 2, \cdots, n$

$$M[x_i] = \int \xi dV_i(\xi) = a_i \qquad \text{(mean value)},$$

$$\text{St. D. } (x_i) = \int (\xi - a_i)^2 dV_i(\xi) = b_i \qquad \text{(standard deviation)},$$

$$\int |\xi - a_i|^3 dV_i(\xi) = c_i \qquad \text{(absolute moment of the 3rd order)}.$$

(The term "standard deviation" is used by some authors for $b_i^{1/2}$ instead of b_i.)

Denoting mean value and standard deviation of the sum $x_1 + \cdots + x_i$ $(i=2, \cdots, n)$ by A_i and B_i respectively, it is known that

$$A_i = a_1 + \cdots + a_i, \qquad B_i = b_1 + \cdots + b_i.$$

For the distribution function W_i of the sum $x_1 + \cdots + x_i$ the recursion formula

$$W_i(x) = \int W_{i-1}(x - \xi) dV_i(\xi), \qquad i > 1, \tag{1}$$

holds true.

Let a_i and B_n be as follows:

$$a_i = 0, \tag{2}$$

$$B_n = 1. \tag{3}$$

These conditions are only preliminarily stated in order to abbreviate the explanation. From (2) we have that $A_i=0$, in particular, $A_n=0$. (The case $b_i=0$ may be excluded, since then x_i assumes the value 0 only with the probability 1 and thus does not appear in the sum $x_1 + \cdots + x_n$.)

[1] Presented to the Society, October 28, 1939.

[2] The following integrals are Stieltjes integrals. The lower and upper limits always are $-\infty$ and $+\infty$ respectively.

We have to prove that the function $W_n(x)$ lies uniformly in x near the Gaussian function having also the mean value 0 and the standard deviation 1, that is,

$$\phi(x/2^{1/2}) = \frac{1}{\pi^{1/2}} \int_{-\infty}^{x/2^{1/2}} e^{-t^2} dt,$$

if very general suppositions with respect to the V are made. We suppose that the quotients c_i/b_i are "very small" for all i and state the general limit theorem in the following form:

To every positive ϵ there is a positive δ, so that uniformly in x

$$| W_n(x) - \phi(x/2^{1/2}) | < \epsilon$$

holds true if the conditions

$$c_i/b_i < \delta, \qquad i = 1, 2, \cdots, n, \tag{4}$$

are satisfied.

Our proof is suggested by that of A. Khintchine (see below). We remark first of all that

$$b_i < \delta, \qquad i = 1, 2, \cdots, n, \tag{5}$$

follows from (4) according to the well known inequality $b_i^3 < c_i^2$. (Of course we can suppose $\delta < 1$.) Since $B_n = 1$ but the b_i are small, it is obvious that n must be a large number.

Now we note that the function of two variables

$$\phi(x/(2y)^{1/2}) = \frac{1}{\pi^{1/2}} \int_{-\infty}^{x/(2y)^{1/2}} e^{-t^2} dt = U(x, y),$$

that is, the Gaussian function with the variable standard deviation y represents a solution of the differential equation of the conduction of heat

$$\frac{\partial \phi}{\partial y} = \frac{1}{2} \frac{\partial^2 \phi}{\partial x^2}. \tag{6}$$

Here and in what follows it is always supposed that $y \geqq \eta$, η being a positive small but constant magnitude. The equation (1) suggests that we compare the two expressions $U(x, B_i)$ and

$$\int U(x - \xi, B_{i-1}) dV_i(\xi).$$

(Later the variable y will be put equal to B_{i-1}.) According to Taylor's

formula we have

$$U(x, y + b_i) = U(x, y) + b_i \frac{\partial U(x, y)}{\partial y} + R_1,$$

$$R_1 = \frac{b_i^2}{2} \frac{\partial^2 U(x, y + \theta b_i)}{\partial y^2}, \qquad 0 < \theta < 1.$$

Let $M(\eta)$ be an upper limit of the absolute values of the derivatives of the 2d and 3d order of $U(x, y)$ in the half-plane $y \geqq \eta$; then

$$|R_1| < b_i^2 M(\eta)/2.$$

On the other hand,

$$U(x - \xi, y) = U(x, y) - \xi \frac{\partial U(x, y)}{\partial x} + \frac{\xi^2}{2} \frac{\partial^2 U(x, y)}{\partial x^2}$$

$$- \frac{\xi^3}{6} \frac{\partial^3 U(x - \theta_1 \xi, y)}{\partial x^3}, \qquad 0 < \theta_1 < 1.$$

Therefore, in view of (2) and (6)

$$\int U(x - \xi, y) dV_i(\xi) = U(x, y) + b_i \frac{\partial U(x, y)}{\partial y} + R_2,$$

where

$$R_2 = -\frac{1}{6} \int \xi^3 \frac{\partial^3 U(x - \theta_1 \xi, y)}{\partial x^3} dV_i(\xi);$$

thus

$$|R_2| < c_i M(\eta)/6.$$

It follows that

$$U(x, y + b_i) = \int U(x - \xi, y) dV_i(\xi) + R \tag{7}$$

in which $R = R_1 - R_2$; accordingly

$$|R| \leqq |R_1| + |R_2| < M(\eta)\{b_i^2/2 + c_i/6\} < [M(\eta)/2]\{b_i^2 + c_i\}. \tag{8}$$

Now note that all $b_i < \eta$ because of (5), since of course $\delta < \eta$ may be supposed. On the contrary there are numbers $B_i > \eta$, as for instance $B_n = 1 > \eta$. There must be consequently a number B_r satisfying $\eta < B_r < 2\eta$. Now we put $y = B_{i-1}$ in (7); and we suppose $i \geqq r+1$ in order that $y \geqq B_r > \eta$. Besides we write with regard to what follows

$x+2\alpha$ instead of x, α representing a positive parameter. Then it follows from (7) that

$$U(x+2\alpha, B_i) = \int U(x+2\alpha-\xi, B_{i-1})dV_i(\xi) + R \tag{9}$$

in which R satisfies (8). Putting the difference

$$W_i(x) - U(x+2\alpha, B_i) = W_i(x) - \phi\left(\frac{x+2\alpha}{(2B_i)^{1/2}}\right) = D_i(x), \tag{10}$$
$$i = 1, 2, \cdots, n,$$

and subtracting (9) from (1), we get the fundamental formula, valid for all x,

$$D_i(x) = \int D_{i-1}(x-\xi)dV_i(\xi) - R \tag{11}$$

together with (8), where $i=r+1, r+2, \cdots, n$.

In order to get further estimates we need the elementary inequalities of Tschebyscheff, which are valid for any distribution function F with vanishing mean value and standard deviation not greater than B:

$$F(\alpha) \geqq 1 - B/\alpha^2, \qquad F(-\alpha) \leqq B/\alpha^2, \qquad \alpha > 0.$$

From these inequalities it easily follows for two distribution functions F_1, F_2 with vanishing mean values, the standard deviations of which are both not greater than B, and for all x, that

$$F_1(x) - F_2(x+2\alpha) \leqq B/\alpha^2, \qquad \alpha > 0. \tag{12}$$

For, if $x \leqq -\alpha$, one has $F_1(x) \leqq F_1(-\alpha) \leqq B/\alpha^2$; thus (12) follows because $F_2 \geqq 0$. If $x > -\alpha$, one concludes $F_2(x+2\alpha) \geqq F_2(\alpha) \geqq 1 - B/\alpha^2$; therefore

$$F_1(x) - F_2(x+2\alpha) \leqq 1 - F_2(x+2\alpha) \leqq 1 - (1 - B/\alpha^2),$$

that is, (12). On the foundation of (11) and (12) the proof may be carried out. Putting $i=r$ one has, according to the definition of $D_i(x)$ and because both $W_r(x)$ and $U(x, B_r)$ have the mean value 0 and the standard deviation $B_r < 2\eta$, on account of (12),

$$W_r(x) - U(x+2\alpha, B_r) = D_r(x) < 2\eta/\alpha^2 \quad \text{for all } x, \qquad \alpha > 0; \tag{13}$$

whence, denoting by g_i the upper limit of $D_i(x)$, we conclude from (13) $g_r < 2\eta/\alpha^2$ and from (11), for any $i \geqq r+1$, $D_i(x) \leqq g_{i-1} + |R|$; thus

$$g_i \leqq g_{i-1} + [M(\eta)/2]\{b_i^2 + c_i\}$$

and through induction from i to $i+1$

$$(14) \quad g_n \leqq g_r + \frac{M(\eta)}{2} \sum_{i=r+1}^{n} \{b_i^2 + c_i\} < \frac{2\eta}{\alpha^2} + \frac{M(\eta)}{2} \sum_{i=1}^{n} \{b_i^2 + c_i\}.$$

On the other hand one has

$$b_i < \delta, \qquad c_i/b_i < \delta, \qquad i = 1, 2, \cdots,$$

because of (5) and (4). Therefore

$$\sum_{i=1}^{n} b_i^2 < \delta \sum b_i = \delta B_n = \delta, \qquad \sum_{i=1}^{n} c_i < \delta \sum b_i = \delta.$$

It follows that $g_n < 2\eta/\alpha^2 + \delta M(\eta)$; thus

$$D_n(x) = W_n(x) - \phi\left(\frac{x+2\alpha}{2^{1/2}}\right) < 2\eta/\alpha^2 + \delta M(\eta).$$

Therefore

$$(15) \qquad W_n(x) < \phi(x/2^{1/2}) + 2\alpha/(2\pi)^{1/2} + 2\eta/\alpha^2 + \delta M(\eta).$$

The positive number ϵ being given, one determines first α through $2\alpha/(2\pi)^{1/2} = \epsilon/3$, then η through $2\eta/\alpha^2 = \epsilon/3$, with which $M(\eta)$ is determined, and finally δ through $\delta M(\eta) \leqq \epsilon/3$. It follows that

$$(16) \qquad W_n(x) < \phi(x/2^{1/2}) + \epsilon,$$

uniformly in x. Using

$$\phi\left(\frac{x-2\alpha}{(2B_i)^{1/2}}\right) - W_i(x) = \Delta_i(x)$$

one finds in the same way

$$\Delta_n(x) < 2\eta/\alpha^2 + \delta M(\eta)$$

and

$$(17) \qquad W_n(x) > \phi(x/2^{1/2}) - \epsilon.$$

The proof is thus complete.

One should compare this with the proof given by A. Khintchine (*Asymptotische Gesetze der Wahrscheinlichkeitsrechnung*, Ergebnisse der Mathematik und ihrer Grenzgebiete, vol. 2, no. 4, Berlin, 1933), which is based on works of I. Petrowsky and A. Kolmogoroff. We have chosen other assumptions in order to make the exposition as simple as possible, while Khintchine strives for wide generality so that nothing is supposed by him as to moments of the third order.

A so-called upper- and lower-function, the introduction of which is denoted by Khintchine as the fundamental idea of the Petrowsky-Kolmogoroffian proof, is not used by us.

We now generalize the investigation and omit the special conditions (2) and (3).

We use the transformation

$$\frac{x_i - a_i}{B_n^{1/2}} = u_i, \qquad i = 1, \cdots, n. \tag{18}$$

If mean value, standard deviation, and absolute moment of the third order of u_i are denoted by $\bar{a}_i$, $\bar{b}_i$, and $\bar{c}_i$ respectively, it becomes

$$\bar{a}_i = 0, \quad \bar{b}_i = b_i/B_n, \quad \bar{c}_i = c_i/(B_n)^{3/2} \tag{2'}$$

and accordingly

$$\frac{\bar{c}_i}{\bar{b}_i} = \frac{c_i}{b_i(B_n)^{1/2}},$$

and

$$\overline{B}_n = \text{St. D. } (u_1 + \cdots + u_n) = \bar{b}_1 + \cdots + \bar{b}_n = 1. \tag{3'}$$

The assumptions (4) must be replaced by

$$\frac{c_i}{b_i B_n^{1/2}} < \delta, \qquad i = 1, \cdots, n. \tag{4'}$$

Denoting by $\overline{W}_n$ the distribution function of the sum $u_1 + \cdots + u_n$, we thus have according to (16), (17) uniformly in u

$$\overline{W}_n(u) = \phi(u/2^{1/2}) + \sigma, \qquad |\sigma| < \epsilon. \tag{19}$$

But it easily follows in view of

$$u_1 + \cdots + u_n = (x_1 + \cdots + x_n - A_n)/B_n^{1/2}$$

that the distribution function $W_n(x)$ of $x_1 + \cdots + x_n$ is identical with $\overline{W}_n(u)$ if the relation

$$u = (x - A_n)/B_n^{1/2} \tag{20}$$

between u and x is valid.

According to (19) and $W_n(x) = \overline{W}_n(u)$ we have uniformly in x

$$W_n(x) = \phi\left(\frac{x - A_n}{(2B_n)^{1/2}}\right) + \sigma, \qquad |\sigma| < \epsilon. \tag{21}$$

We can thus say: *To every positive ϵ there is a positive δ, so that* (21) *holds true, if the assumptions* (4′) *are satisfied.*

We call special attention to the case

$$0 < g < b_i; \qquad c_i < H,$$

g and H being independent of i. One has $gn < B_n$ and as for the rest $b_i < c_i^{2/3} < H^{2/3}$ and $B_n < nH^{2/3}$, so that B_n is exactly of the order n. It follows that

$$\frac{c_i}{b_i B_n^{1/2}} < \frac{H}{g g^{1/2} n^{1/2}}.$$

Therefore (4′) can be satisfied, if n is sufficiently large.

In the same way one can discuss sets of chance variables $x_{11}, x_{12}, \cdots, x_{1n_1}, \cdots, x_{k1}, x_{k2}, \cdots, x_{kn_k}, \cdots, n_k$ increasing indefinitely with k. It is supposed that the quotients

$$\frac{c_{k1}}{b_{k1} B_{n_k}^{1/2}}, \cdots, \frac{c_{kn_k}}{b_{kn_k} B_{n_k}^{1/2}}$$

approach 0 as a limit, if k increases indefinitely. That means that if any $\delta > 0$ is given one can choose k sufficiently large so that *all* these quotients are less than δ. Then the distribution function W_{n_k} of $x_{k1} + \cdots + x_{kn_k}$ approaches $\phi((x - A_{n_k})/(2B_{n_k})^{1/2})$ as a limit.

The conditions for the quotients are in particular satisfied if $B_{n_k} > B > 0$ for every k (B being a constant number) and the quotients $c_{k1}/b_{k1}, \cdots, c_{kn_k}/b_{kn_k}$ approach 0, as k increases indefinitely.

NEW YORK, N. Y.

A NEW LOWER BOUND FOR THE EXPONENT IN THE FIRST CASE OF FERMAT'S LAST THEOREM[1]

BARKLEY ROSSER

1. **Introduction.** In this paper is proved the theorem: If p is an odd prime and

$$a^p + b^p + c^p = 0 \tag{1}$$

has a solution in integers prime to p, then $p > 41{,}000{,}000$.

It seems certain that still higher lower bounds for p can be deduced by the methods of this paper. However an argument is given which makes it seem unlikely that an indefinitely high lower bound can be so deduced.

2. **Preliminary results.** Unless otherwise specified, we shall assume that p is an odd prime for which (1) can be satisfied by integers prime to p. Hence[2] $p > 8{,}000{,}000$. Also $x \equiv y$ shall denote $x \equiv y \pmod{p}$. Also any statement regarding factorization of a polynomial is to be understood modulo p.

Morishima has proved[3] that for each odd prime $m \leqq 43$, there is a t ($t \not\equiv 0, \not\equiv 1$) such that each of the values

$$t, \quad \frac{1}{t}, \quad 1-t, \quad \frac{1}{1-t}, \quad \frac{t-1}{t}, \quad \frac{t}{t-1} \tag{2}$$

satisfies each of the following relations when substituted for x:

$$\{(m^{p-1} - 1)/p\}(x^{m-1} - 1) \equiv 0, \tag{3}$$

$$x^4 \not\equiv 1, \tag{4}$$

$$x^6 \not\equiv 1. \tag{5}$$

Morishima further proved that for each odd prime $m \leqq 31$, there is no t such that the values in (2) satisfy (4), (5), and

$$x^{m-1} \equiv 1. \tag{6}$$

Hence for such m's, $(m^{p-1} - 1)/p \equiv 0$. That is

$$m^{p-1} \equiv 1 \pmod{p^2}. \tag{7}$$

---

[1] Presented to the Society, September 8, 1939.

[2] Barkley Rosser, *On the first case of Fermat's last theorem*, this Bulletin, vol. 45 (1939), pp. 636–640. This paper will be referred to as I.

[3] Taro Morishima, *Über den Fermatschen Quotienten*, Japanese Journal of Mathematics, vol. 8 (1931), pp. 159–173.

In this paper we shall use the fact that $p>8{,}000{,}000$ to prove that, for $m=37$ or 41, there is no t such that the values in (2) satisfy (4), (5), and (6). Hence we can conclude that (7) holds for each prime $m\leqq 41$ (since it is well known that (7) holds for $m=2$). This fact is then used to prove that $p>41{,}000{,}000$.

Until §6, we assume that m is an odd prime, $n=(m-1)/2$, and that a t exists such that the values of (2) satisfy (4), (5), and (6).

We reduce the consideration of (4), (5), and (6) to a familiar theory[4] by replacing t by $-a$. Then since the polynomials in (4), (5), and (6) are all even functions, we see that each of

$$\begin{aligned} &a,\ \frac{1}{a},\ -1-a,\ \frac{1}{-1-a},\ \frac{1+a}{-a},\ \frac{-a}{1+a},\\ &-a,\ \frac{1}{-a},\ 1+a,\ \frac{1}{1+a},\ \frac{1+a}{a},\ \frac{a}{1+a}, \end{aligned} \tag{8}$$

must satisfy (4), (5), and (6). Put

$$b=\frac{-a^6-3a^5+5a^3-3a-1}{a^2(a+1)^2},$$

and define

$$f(x)=(x-a)(x-1/a)(x+1+a)(x+1/(1+a))(x+(1+a)/a)(x+a/(1+a)).$$

Then $f(x)=x^6+3x^5+bx^4+(2b-5)x^3+bx^2+3x+1$.

We shall prove in §3 that $f(x)f(-x)$ has no multiple factors if $m\leqq 67$. So for $m\leqq 67$, $f(x)f(-x)$ must divide $x^{m-1}-1$. Also note that $f(x)f(-x)$ has no factors in common with either x^4-1 or x^6-1, by (4) and (5).

LEMMA 1. *If $\phi(x)$ is a polynomial in x and $f(x)f(-x)$ has no multiple factors, then a n.a.s.c. that $f(x)f(-x)$ divide $\phi(x)$ is that $f(x)$ divide both $\phi(x)$ and $\phi(x+1)$.*

PROOF. By comparing the values in (8), one quickly sees that $f(-x)=f(x-1)$. So $f(-x)$ divides $\phi(x)$ if and only if $f(x-1)$ divides $\phi(x)$, if and only if $f(x)$ divides $\phi(x+1)$.

LEMMA 2. *If $f(x)f(-x)$ has no multiple factors and divides $x^{4N}-1$, then either*

A. *$f(x)$, $x^{2N}-1$, and $(x+1)^{2N}+1$ have a common factor, or*

B. *$f(x)f(-x)$ divides $x^{2N}-1$.*

[4] In particular we shall use the results of L. E. Dickson, *On the last theorem of Fermat*, Messenger of Mathematics, vol. 38 (1908), pp. 14–32.

PROOF. Use Lemma 1 and the results of Dickson[4] (loc. cit., §11, pp. 20–21).

3. **$f(x)f(-x)$ has no multiple factors if $m \leqq 67$.** By (4) and (5) one can readily see that the only cases in which $f(x)f(-x)$ can have a multiple root are:

I. $a^2+a-1\equiv 0$,
II. $a^2-a-1\equiv 0$,
III. $a^2+3a+1\equiv 0$.

In case I, $x^{2n}-1$ and x^2-x-1 have the factor $x+a$ in common by (6). In case II, $x^{2n}-1$ and x^2-x-1 have the factor $x-a$ in common by (6). In case III, $x^{2n}-1$ and x^2-x-1 have the factor $x+a+1$ in common by (6). So it suffices to prove that $x^{2n}-1$ and x^2-x-1 have no factors in common.

LEMMA 3. *If $n \leqq 33$, then $x^{2n}-1$ and x^2-x-1 have no common factor modulo p.*

PROOF. Divide $x^n \pm 1$ by x^2-x-1. If $a_1, a_2, \cdots$ are the Fibonacci numbers, that is $a_1=a_2=1$, $a_{j+2}=a_{j+1}+a_j$, and $Q_j(x)=a_1x^j+a_2x^{j-1}+\cdots+a_jx+a_{j+1}$, then $x^n\pm 1=(x^2-x-1)Q_{n-2}(x)+a_nx+(a_{n-1}\pm 1)$. The eliminant of x^2-x-1 and $a_nx+(a_{n-1}\pm 1)$ is $(a_{n-1})^2+a_{n-1}a_n-(a_n)^2\pm(2a_{n-1}+a_n)+1$. However $(a_{j-1})^2+a_ja_{j-1}-(a_j)^2=(-1)^j$. So the eliminant of $x^n\pm 1$ and x^2-x-1 is $1+(-1)^n\pm(2a_{n-1}+a_n)$. However $a_j=(k^j-l^j)/5^{1/2}$, where $k=(1+5^{1/2})/2$, $l=(1-5^{1/2})/2$. So a_j equals the integer nearest to $k^j/5^{1/2}$. Hence $x^n\pm 1$ and x^2-x-1 have no common factor if $n\leqq 33$, since $p>8{,}000{,}000$ and factorization is modulo p. Therefore $x^{2n}-1$ and x^2-x-1 have no common factor if $n\leqq 33$.

4. **Proof that $m\neq 37$.** Assume $m=37$, and then use Lemma 2 with $N=9$.

Case 1. $x-\alpha$ is a common factor of $f(x)$, $x^{18}-1$, and $(x+1)^{18}+1$. By (5), $x-\alpha$ is not a factor of x^6-1. So $x-\alpha$ is a common factor of $(x^6+x^3+1)(x^6-x^3+1)$ and $(x+1)^{18}+1$.

Subcase I. $x-\alpha$ is a common factor of x^6+x^3+1 and $(x+1)^{18}+1$. So $\alpha^9\equiv 1$ and $(\alpha^2+2\alpha+1)^9/\alpha^9+1\equiv 0$. Put $\beta=\alpha+1/\alpha$. Then $A=\beta^3-3\beta+1\equiv 0$, and, since $\beta+2=(\alpha+1)^2/\alpha$,

$$(\beta+2)^9+1=((\beta+2)^3+1)((\beta+2)^6-(\beta+2)^3+1)\equiv 0.$$

If $(\beta+2)^3+1\equiv 0$, then $6\beta^2+15\beta+8\equiv 0$, $23\beta+52\equiv 0$, and[5] $2516=2^2\cdot 17\cdot 37\equiv 0$. So

[5] Here, and at corresponding places later, a contradiction results from the fact that $p>8{,}000{,}000$.

$$(\beta+2)^6-(\beta+2)^3+1\equiv 0, \quad B=417\beta^2+699\beta-137\equiv 0,$$
$$B\beta-417A=C=699\beta^2+1096\beta-411\equiv 0,$$
$$3B-C=534\beta^2+1001\beta\equiv 0.$$

However $\beta\equiv 0$ would contradict $A\equiv 0$. So $534\beta+1001\equiv 0$. So $883{,}227=3\cdot 37\cdot 73\cdot 109\equiv 0$.

Subcase II. $x-\alpha$ is a common factor of x^6-x^3+1 and $(x+1)^{18}+1$. So $\alpha^9\equiv -1$ and $(\alpha^2+2\alpha+1)^9/\alpha^9\equiv 1$. Put $\beta=\alpha+1/\alpha$. Then

$$\beta^3-3\beta-1\equiv 0,$$
$$(\beta+2)^9-1\equiv((\beta+2)^3-1)((\beta+2)^6+(\beta+2)^3+1)\equiv 0.$$

If $(\beta+2)^3-1\equiv 0$, then $6\beta^2+15\beta+8=0$, $23\beta+28\equiv 0$, and $724=2^2\cdot 181\equiv 0$. So

$$(\beta+2)^6+(\beta+2)^3+1\equiv 0, \quad 447\beta^2+861\beta+271\equiv 0,$$
$$37(187\beta+302)\equiv 0.$$

Using $187\beta+302\equiv 0$, we get $1{,}620{,}673=73\cdot 149^2\equiv 0$.

Case 2. $f(x)f(-x)$ divides $x^{18}-1$. Because of (5), $f(x)f(-x)$ divides $x^{12}+x^6+1$. So $f(x)f(-x)\equiv x^{12}+x^6+1$. Comparing the coefficients of x^{10} and x^8, we get $2b-9\equiv 0$ and $b^2-10b+30\equiv 0$. So $21\equiv 0$.

5. **Proof that** $m\neq 41$. Assume $m=41$. Use Lemma 2 with $N=10$.

Case 1. $f(x)$, $x^{20}-1$, and $(x+1)^{20}+1$ have a common factor. The discussion in Dickson (loc. cit., pp. 23–24) will eliminate this case.

Case 2. $f(x)f(-x)$ is a factor of $x^{20}-1$. By Lemma 2 with $N=5$, either $f(x)$, $x^{10}-1$ and $(x+1)^{10}+1$ have a common factor, or else $f(x)f(-x)$ divides $x^{10}-1$. The latter is impossible, so let $x-\alpha$ be a common factor of $f(x)$, $x^{10}-1$, and $(x+1)^{10}+1$. Because of (4), $x-\alpha$ is not a factor of x^2-1.

Subcase I. $x-\alpha$ is a common factor of $x^4+x^3+x^2+x+1$ and $(x+1)^{10}+1$. Then $\alpha^5\equiv 1$. So $(\alpha^2+2\alpha+1)^5/\alpha^5+1\equiv 0$. Put $\beta=\alpha+1/\alpha$. Then $\beta^2\equiv -\beta+1$, $(\beta+2)^5+1\equiv 0$. So $5(11\beta+16)\equiv 0$. So $41\equiv 0$.

Subcase II. $x-\alpha$ is a common factor of $x^4-x^3+x^2-x+1$ and $(x+1)^{10}+1$. Then $\alpha^5\equiv -1$. So put $\beta=\alpha+1/\alpha$. Then $\beta^2\equiv\beta+1$, $(\beta+2)^5-1\equiv 0$. So $275\beta+174\equiv 0$. So $2501=41\cdot 61\equiv 0$.

6. **Proof that** $p>41{,}000{,}000$. We have now proved

$$37^{p-1}\equiv 1 \pmod{p^2}, \quad 41^{p-1}\equiv 1 \pmod{p^2}.$$

So we can argue as in I and conclude that $p>41{,}000{,}000$.

7. **The non-generality of this method.** It seems certain that the

lower bound just obtained is not the best that this method will produce. For instance, now that we know that $p>41{,}000{,}000$, it is almost certain that we can prove that (7) holds for $m=43$ (in fact, this probably follows from $p>8{,}000{,}000$). Also, the fact that $p>41{,}000{,}000$ will probably make it easier to prove that, for appropriate $m>43$, the values of (2) must satisfy (3), (4), and (5), and so, for these m's also, (7) undoubtedly holds. Nevertheless the methods of this paper will apparently fail for very large values of p.

To see the difficulty that would arise, let us ignore any difficulties inherent in the proof that the values of (2) must satisfy (3), (4), and (5), and assume that for any m, this result would be easily forthcoming. Then, for any m, (7) would hold except for these p's which are factors of the eliminant of $x^{m-1}-1$ and $(x+1)^{m-1}-1$. As $p>41{,}000{,}000$, we could then prove that (7) holds for all $m\leqq N_1$ (undoubtedly $N_1>43$). From this, by the method of I, we could prove that $p>M_1$. From this we could then prove that (7) holds for all $m\leqq N_2$, and so on. Unfortunately the M's and N's would probably not increase indefinitely.

To see this, let us first consider the eliminant of $x^{m-1}-1$ and $(x+1)^{m-1}-1$. Let $x-\alpha$ be a common factor. Then $\alpha^{m-1}\equiv 1$ and $(\alpha+1)^{m-1}\equiv 1$. We break this into cases (essentially corresponding to a factorization of the eliminant (see Dickson,[6] p. 28), by considering the eliminants of $(x+1)^u-x^v$ and $x^{m-1}-1$ for small values of u and v. The case $u=1$, $v=2$ was treated in Lemma 3, and two eliminants of the order $k^{m/2}$ were obtained. If m is the rth prime, m is of the order of $r\log r$ (see Landau,[7] pp. 213–215). So the eliminants are of the order of r^{br}.

Now we consider the lower bounds for p which can be obtained by the method of I. Let the Σ's be as in Lemma 5 of I and let r be large. Then $\Sigma_1=\theta(p_r)-\log 2<2r\log r$ (Landau, loc. cit., p. 195). So $\Sigma_j<(\Sigma_1)^j/j!<(2r\log r)^j/j!$. Also, since r is large, $r!(\log 2)\Sigma_{r-1}>(e^{-r}r^r)e^r=r^r$. Substituting these in the formula for $f^r(x)$ in Lemma 5 of I, we get

$$f_r(x)<(x+r\log r)^r/r^r.$$

If $x>(2\log r)^r$, then $f_r(\log x^2/2)<x/2$. So the method of I will not give a lower bound as great as $(2\log r)^r$, whereas the eliminant of x^{p_r-1} and x^2-x-1 is of the order of r^{br}.

To the above argument, one might object that firstly it is not the

[6] L. E. Dickson, *On the last theorem of Fermat*, Quarterly Journal of Mathematics, vol. 40 (1908), pp. 27–45.

[7] E. Landau, *Handbuch der Lehre von der Verteilung der Primzahlen.*

size of the eliminants but the size of their largest prime factor which is important, and secondly it is not essential to take the m's in order of magnitude. In answer, it should be pointed out that after one passes the limits of factor tables, it becomes impracticable to deal with the factors of the eliminant rather than the eliminant. Therefore, since the eliminant (in one case at least) appears to be an increasing function of m, one is compelled to work with monotone increasing m.

CORNELL UNIVERSITY

SOME UNIFORMLY CONVEX SPACES

R. P. BOAS, JR.[1]

1. **Introduction.** A Banach space is said to be uniformly convex if to every ϵ, $0<\epsilon<1$, there is a $\delta(\epsilon)$, $0<\delta(\epsilon)<2$, such that $\|x\|=\|y\|=1$ and $\|x-y\|\geqq\epsilon$ imply $\|x+y\|<2-\delta(\epsilon)$. J. A. Clarkson, who introduced the concept of uniform convexity [5], proved that the spaces L^p and l^p are uniformly convex if $p>1$, basing his proof on the following inequalities[2] among norms of elements of L^p or l^p:

$$(1.1)\qquad \|x+y\|^p+\|x-y\|^p\leqq 2^{p-1}(\|x\|^p+\|y\|^p), \qquad p\geqq 2;$$

$$(1.2)\qquad \|x+y\|^p+\|x-y\|^p\leqq 2(\|x\|^{p'}+\|y\|^{p'})^{p-1}, \qquad p\geqq 2;$$

$$(1.3)\qquad \|x+y\|^{p'}+\|x-y\|^{p'}\leqq 2(\|x\|^p+\|y\|^p)^{p'-1}, \quad 1<p\leqq 2.$$

The uniform convexity of L^p and l^p follows by inspection from either (1.1) or (1.2) if $p\geqq 2$, and from (1.3) if $1<p\leqq 2$. As Clarkson observed, (1.1) is a consequence of (1.2), since $\{(1/2)(a^r+b^r)\}^{1/r}$ is an increasing function of r for positive a and b [6, p. 26], so that the right side of (1.1) is not less than that of (1.2). However, (1.1) is interesting because it is considerably simpler to prove than (1.2) (see §3), so that the uniform convexity of L^p and l^p can be established more easily for $p\geqq 2$ than for $1<p<2$.

In this note I give a short proof of Clarkson's inequalities (and of a general set of inequalities, which includes them), using M. Riesz's convexity theorem for linear forms. This proof has the advantage that it can be generalized to show that the spaces $L^p\{L^q\}$, $L^p\{l^q\}$, $l^p\{L^q\}$, $l^p\{l^q\}$ are all uniformly convex[3] if $p>1, q>1$. Here $L^p\{E\}$ is

[1] National Research Fellow.

[2] Here, as throughout this note, $p'=p/(p-1)$; similarly for other letters.

[3] These results suggest the possibility that $L^p\{E\}$ and $l^p\{E\}$ are uniformly convex whenever E is; but I can offer no evidence for or against this conjecture.

the Banach space of functions $x(t)$ from a real interval to the Banach space E, integrable (in the sense of Bochner [3]), and such that $\mathfrak{M}_p[x]=(\int\|x(t)\|^p dt)^{1/p}$ is finite; the norm of x in $L^p\{E\}$ is[4] $\mathfrak{M}_p[x]$. Similarly, $l^p\{E\}$ is the Banach space of sequences $X=\{x_n\}$ of elements of E, such that $\mathfrak{M}_p[X]=(\sum\|x_n\|^p)^{1/p}$ is finite, with $\mathfrak{M}_p[X]$ as the norm of X.

I also give a separate proof of (1.1); this proof is strictly elementary, depending only on the inequalities of Hölder, Minkowski, and Jensen.

Clarkson's inequalities can be written in the general form

$$(1.4)\qquad (\|x+y\|^r+\|x-y\|^r)^{1/r}\leqq 2^{1/s'}(\|x\|^s+\|y\|^s)^{1/s};$$

(1.4) reduces to (1.1), (1.2), (1.3) for $r=s=p$; $r=p$, $s=p'$; $r=p'$, $s=p$. In each case, $s\leqq p\leqq r$; this condition will be used in §2.

2. **A preliminary reduction.** We begin by showing that (1.4), for an r and an s with $1<s\leqq p\leqq r$, is implied by the same inequality for complex x and y (with norms replaced by absolute values). In fact, if we know

$$(2.1)\qquad (|x+y|^r+|x-y|^r)^{1/r}\leqq 2^{1/s'}(|x|^s+|y|^s)^{1/s}$$

for all complex x and y, with $1<s\leqq p\leqq r$, we deduce (1.4) (for L^p) from the inequalities [6, p. 148]

$$(2.2)\qquad \left\{\left(\int|x(t)+y(t)|^p dt\right)^{r/p}+\left(\int|x(t)-y(t)|^p dt\right)^{r/p}\right\}^{1/r}$$
$$\leqq\left\{\int\left(|x(t)+y(t)|^r+|x(t)-y(t)|^r\right)^{p/r}dt\right\}^{1/p},\quad r\geqq p;$$

$$(2.3)\qquad 2^{1/s'}\left\{\int\left(|x(t)|^s+|y(t)|^s\right)^{p/s}dt\right\}^{1/p}$$
$$\leqq 2^{1/s'}\left\{\left(\int|x(t)|^p dt\right)^{s/p}+\left(\int|y(t)|^p dt\right)^{s/p}\right\}^{1/s},\quad s\leqq p.$$

For, (2.1) shows that the left side of (2.3) is not less than the right side of (2.2). For l^p, we have the same inequalities with $\int$ replaced by $\sum$ [6, p. 123], and we draw the same conclusions.

3. **Independent proof of** (1.1). According to §2, it is enough to prove

$$(3.1)\qquad (|x+y|^p+|x-y|^p)^{1/p}\leqq 2^{1/p'}(|x|^p+|y|^p)^{1/p},\quad p\geqq 2.$$

[4] See Bochner and Taylor [4] for a discussion of these spaces.

We have, using first Jensen's inequality [6, p. 28] and then Hölder's inequality with index $p/2$ [6, p. 24],

$$\begin{aligned}(|x+y|^p+|x-y|^p)^{1/p} &\leqq (|x+y|^2+|x-y|^2)^{1/2}\\ &= 2^{1/2}(|x|^2+|y|^2)^{1/2}\\ &\leqq 2^{1/2}\{2^{(p-2)/p}(|x|^p+|y|^p)^{2/p}\}^{1/2}\\ &= 2^{1/p'}(|x|^p+|y|^p)^{1/p},\end{aligned}$$

since $(p/2)'=p/(p-2)$.

This is the simplest proof of the uniform convexity of L^p and l^p, $(p>2)$.

4. **Proof of Clarkson's inequalities.** We shall prove the following general theorem, of which Clarkson's inequalities are special cases: it asserts the truth of (1.2) and (1.3), respectively, when $p\geqq 2$, $r=p$, $s=p'$, and when $1<p\leqq 2$, $r=p'$, $s=p$.

THEOREM 1. *If x and y are elements of L^p or l^p, $(p>1)$, then*

$$(4.1)\qquad (\|x+y\|^r+\|x-y\|^r)^{1/r}\leqq 2^{1/s'}(\|x\|^s+\|y\|^s)^{1/s}$$

whenever the point (r, s) is in the common part of the regions of the (r, s) plane determined by the inequalities

$$(4.2)\qquad r\geqq s\geqq r',\qquad r\geqq p\geqq s>1.$$

As we saw in §2, it is enough (since the second inequality (4.2) is satisfied) to establish

$$(4.3)\qquad (|x+y|^r+|x-y|^r)^{1/r}\leqq 2^{1/s'}(|x|^s+|y|^s)^{1/s}$$

for all complex x and y. We now appeal to M. Riesz's convexity theorem,[5] which, stated for the pair of linear forms $x+y$ and $x-y$, is that

$$M_{a,b}=\log\max\frac{(|x+y|^{1/a}+|x-y|^{1/a})^a}{(|x|^{1/b}+|y|^{1/b})^b}$$

is convex in the triangle[6] $0\leqq a\leqq b\leqq 1$.

Now, we have the identity

$$(|x+y|^2+|x-y|^2)^{1/2}=2^{1/2}(|x|^2+|y|^2)^{1/2},$$

and the trivial inequalities

[5] [6, p. 219]; a new and interesting proof (and generalization) has been given by Thorin [8].

[6] The maximum is taken over all x and y. When a or b is zero, the numerator or denominator on the right is to be replaced by its limiting value as a or b approaches zero, namely $\max(|x+y|, |x-y|)$ or $\max(|x|, |y|)$.

$$\max(|x+y|, |x-y|) \leqq |x|+|y|,$$
$$\max(|x+y|, |x-y|) \leqq 2\max(|x|, |y|);$$

it follows easily that

$$M_{1/2,1/2} = (1/2)\log 2, \qquad M_{0,1} = 0, \qquad M_{0,0} = \log 2. \tag{4.4}$$

The graph of a continuous convex function lies below any chord. Since the linear function of a and b taking the values log 2, 0, $(1/2)$ log 2 at the points (0, 0), (0, 1), and (1/2, 1/2) is $(1-b)$ log 2, we obtain from (4.4) (writing $a=1/r$, $b=1/s$), the inequality (4.3), valid for any r and s such that $(1/r, 1/s)$ is in the triangle whose vertices are (0, 0), (0, 1), and (1/2, 1/2). These values of r and s are those satisfying $r>1$ and the first inequality (4.2).

The uniform convexity of L^p and l^p is a consequence of any case of Theorem 1; Clarkson's inequalities (1.2) and (1.3) are the strongest cases involving p in a simple way.[7] When $r=s=p\geqq 2$, we have the weakest inequality; we should therefore expect to be able to give the simplest proof in this case.

5. **The uniform convexity of $\lambda^p\{\lambda^q\}$.** It is convenient to have a shorthand notation for our L^p and l^p spaces. We shall use λ to stand for either L or l; a statement involving λ is to be interpreted as the set of statements obtained by the reading of L or l for λ in all possible ways.

The space $\lambda^p\{E\}$ is evidently uniformly convex if its elements satisfy any inequality of the form

$$\{\mathfrak{M}_p^r[x+y] + \mathfrak{M}_p^r[x-y]\}^{1/r} \leqq 2^{1/s'}\{\mathfrak{M}_p^s[x] + \mathfrak{M}_p^s[y]\}^{1/s} \tag{5.1}$$

with $r>1$, $s>1$.

The content of (5.1) will perhaps be clearer if we write it out explicitly for one of its cases. For $L^p\{l^q\}$, it states that for any sequences $\{x_n(t)\}$, $\{y_n(t)\}$ of functions belonging to L^p,

$$\left\{\left[\int (\sum |x_n(t)+y_n(t)|^q)^{p/q}dt\right]^{r/p} + \left[\int (\sum |x_n(t)-y_n(t)|^q)^{p/q}dt\right]^{r/p}\right\}^{1/r}$$
$$\leqq 2^{1/s'}\left\{\left[\int (\sum |x_n(t)|^q)^{p/q}dt\right]^{s/p} + \left[\int (\sum |y_n(t)|^q)^{p/q}dt\right]^{s/p}\right\}^{1/s},$$

[7] The left side of (2.1) is an increasing function of $1/r$; the right side is a decreasing function of $1/s$. Hence we get a better inequality by increasing $1/r$ or $1/s$.

provided that the right side is finite. For $l^p\{L^q\}$, (5.1) is the same inequality with $\int$ and $\sum$ interchanged.

Even when E is λ^q, (5.1) is not true (naturally enough) for all choices of r and s. Reasonable generalizations of Clarkson's inequalities are the cases of (5.1) described by the following table, when E is λ^q.

q	p	r	s
$1<q\leqq 2$	$p\geqq q'$	p	p'
$1<q\leqq 2$	$q\leqq p\leqq q'$	q'	q
$1<q\leqq 2$	$1<p\leqq q$	p'	p
$q\geqq 2$	$p\geqq q$	p	p'
$q\geqq 2$	$q'\leqq p\leqq q$	q	q'
$q\geqq 2$	$1<p\leqq q'$	p'	p

These inequalities are included in the following theorem.

THEOREM 2. *If x and y are elements of $\lambda^p\{\lambda^q\}$, $(p>1, q>1)$, then* (5.1) *is true whenever the point (r, s) is in the common part of the regions determined by the inequalities*

$$r\geqq p\geqq s>1, \qquad r(q-1)\geqq s\geqq r' \tag{5.2}$$

when $1<q\leqq 2$, or

$$r\geqq p\geqq s>1, \qquad r(q'-1)\geqq s\geqq r' \tag{5.3}$$

when $q\geqq 2$.

As we have already seen, any case of Theorem 2 implies that the spaces $\lambda^p\{\lambda^q\}$ are uniformly convex for $p>1$, $q>1$.

To prove Theorem 2, we observe that (5.1) is (1.4) with $\mathfrak{M}_p[\cdots]$ for $\|\cdots\|$. If we replace $|\cdots|$ by $\|\cdots\|$ in (2.2) and (2.3), the norm being the norm in λ^q, (2.2) and (2.3) are of course still valid, and we deduce that (5.1) holds if

$$(\|x+y\|^r+\|x-y\|^r)^{1/r}\leqq 2^{1/s'}(\|x\|^s+\|y\|^s)^{1/s} \tag{5.4}$$

with $1<s\leqq p\leqq r$, the norm being a λ^q norm. We shall prove (5.4) in the cases specified by (5.2) and (5.3), and thus establish the corresponding cases of (5.1).

To do this, we generalize the argument of §4; here Clarkson's inequalities will replace the equality

$$\|x+y\|^2+\|x-y\|^2=2(\|x\|^2+\|y\|^2),$$

which is valid only in euclidean spaces (Jordan and von Neumann [7]).

The form of Riesz's theorem used in §4 is equivalent to another theorem of Riesz on bilinear forms; S. Bochner and the author [2] have observed that the latter remains true, with an appropriate interpretation, when the variables involved are elements of a Banach space. It follows that the convexity theorem for linear forms is also true (see §6). For our special case it states that

$$M_{a,b} = \log \sup \frac{(\|x+y\|^{1/a} + \|x-y\|^{1/a})^a}{(\|x\|^{1/b} + \|y\|^{1/b})^b}$$

is convex in the triangle $0 \leqq a \leqq b \leqq 1$. (The least upper bound is taken as x and y vary over λ^q; the norms, of course, are λ^q norms.)

We may now formulate (5.4) as

$$M_{1/r,1/s} \leqq (1/s') \log 2, \qquad 1 < s \leqq p \leqq r < \infty. \tag{5.5}$$

We have, in λ^q (and in fact in any Banach space),

$$M_{0,0} = \log 2, \qquad M_{0,1} = 0.$$

Clarkson's inequalities (1.2) and (1.3) give us

$$M_{1/q,1/q'} \leqq (1/q) \log 2,\ q \geqq 2; \quad M_{1/q',1/q} \leqq (1/q') \log 2,\ 1 < q \leqq 2.$$

Consequently (5.4) is true whenever $(1/r, 1/s)$ is in the triangle with vertices (0, 0), (0, 1), $(1/q, 1/q')$ if $q \geqq 2$, or in the triangle with vertices (0, 0), (0, 1), $(1/q', 1/q)$ if $1 < q \leqq 2$. These points correspond to points (r, s), with $r > 1$, satisfying the second inequalities (5.2) or (5.3). This completes the proof of Theorem 2.

It should be mentioned that certain cases of (5.2), adequate to establish the uniform convexity of $\lambda^p\{\lambda^q\}$ for some (but not all) p and q, follow directly from Clarkson's inequalities without any convexity theorem. For example, if $q \geqq 2$, (5.4) becomes (1.2) for $r = q$, $s = q'$; hence (5.4) still holds for $r \geqq q$, $s \geqq q'$; therefore, if $p \geqq q'$, we can take $r = s = p$ and thus establish the uniform convexity of $\lambda^p\{\lambda^q\}$ for $p \geqq q'$, $q \geqq 2$.

6. **The abstract convexity theorem.** Let E be a Banach space, with elements x; let its conjugate space have elements f (linear functionals on E). Let $N_{c,b}$ be the least upper bound of

$$\left| \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij} f_j(x_i) \right| \tag{6.1}$$

for

$$\sum_{i=1}^{m} \|x_i\|^{1/b} \leqq 1, \qquad \sum_{j=1}^{n} \|f_j\|^{1/c} \leqq 1. \tag{6.2}$$

It is known[8] that $\log N_{c,b}$ is convex in the triangle $0 \leqq b \leqq 1$, $0 \leqq c \leqq 1$, $b+c \geqq 1$.

We wish to show that if $M_{a,b}$ is the least upper bound of the expression

$$\left\{ \sum_{j=1}^{n} \left\| \sum_{i=1}^{m} a_{ij} x_i \right\|^{1/a} \right\}^{a} \tag{6.3}$$

for

$$\sum_{i=1}^{m} \|x_i\|^{1/b} \leqq 1, \tag{6.4}$$

then $\log M_{a,b}$ is convex in the triangle $0 \leqq a \leqq b \leqq 1$. This we do by showing that $M_{a,b} = N_{c,b}$ if $c = 1-a$.

Let $\{f_j\}$ and $\{x_i\}$ be arbitrary sets satisfying (6.2) with $c = 1-a$. Then, by Hölder's inequality,

$$\begin{aligned} \left| \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij} f_j(x_i) \right| &= \left| \sum_{j=1}^{n} f_j\left(\sum_{i=1}^{m} a_{ij} x_i \right) \right| \\ &\leqq \left(\sum_{j=1}^{n} \|f_j\|^{1/(1-a)} \right)^{1-a} \left(\sum_{j=1}^{n} \left\| \sum_{i=1}^{m} a_{ij} x_i \right\|^{1/a} \right)^{a} \\ &\leqq M_{a,b}. \end{aligned}$$

Consequently[9]

$$N_{c,b} \leqq M_{a,b}. \tag{6.5}$$

Now let $\{x_i\}$ be an arbitrary set satisfying (6.4). We can find[10] linear functionals f_j, $(j = 1, 2, \cdots, n)$, such that

$$f_j\left(\sum_{i=1}^{m} a_{ij} x_i \right) = \left\| \sum_{i=1}^{m} a_{ij} x_i \right\| \cdot \|f_j\|, \qquad j = 1, 2, \cdots, n;$$

$$\|f_j\|^{1/c} = A \left\| \sum_{i=1}^{m} a_{ij} x_i \right\|^{1/a}, \qquad j = 1, 2, \cdots, n,$$

with constant $A \neq 0$; and $\sum_{j=1}^{n} \|f_j\|^{1/c} = 1$. For these f_j we have, using the conditions for equality in Hölder's inequality,

[8] Boas and Bochner [2, p. 64].

[9] This simple argument was suggested by the referee.

[10] Banach [1, p. 55].

$$\sum_{j=1}^{n}\sum_{i=1}^{m} a_{ij}f_j(x_i) = \sum_{j=1}^{n} f_j\left(\sum_{i=1}^{m} a_{ij}x_i\right)$$

$$= \sum_{j=1}^{n} \|f_j\| \cdot \left\| \sum_{i=1}^{m} a_{ij}x_i \right\|$$

$$= \left(\sum_{j=1}^{n} \|f_j\|^{1/(1-a)}\right)^{1-a} \left(\sum_{j=1}^{n} \left\| \sum_{i=1}^{m} a_{ij}x_i \right\|^{1/a}\right)^{a}$$

$$= \left(\sum_{j=1}^{n} \left\| \sum_{i=1}^{m} a_{ij}x_i \right\|^{1/a}\right)^{a}.$$

Hence $M_{a,b} \leqq N_{c,b}$ if $c=1-a$; this, with (6.5), completes the proof.

References

1. S. Banach, *Théorie des Opérations Linéaires*, 1932.
2. R. P. Boas, Jr., and S. Bochner, *On a theorem of M. Riesz for Fourier series*, Journal of the London Mathematical Society, vol. 14 (1939), pp. 62–73.
3. S. Bochner, *Integration von Funktionen, deren Werte die Elemente eines Vektorraumes sind*, Fundamenta Mathematicae, vol. 20 (1933), pp. 262–276.
4. S. Bochner and A. E. Taylor, *Linear functionals on certain spaces of abstractly-valued functions*, Annals of Mathematics, (2), vol. 39 (1938), pp. 913–944.
5. J. A. Clarkson, *Uniformly convex spaces*, Transactions of this Society, vol. 40 (1936), pp. 396–414.
6. G. H. Hardy, J. E. Littlewood, and G. Pólya, *Inequalities*, 1934.
7. P. Jordan and J. von Neumann, *On inner products in linear, metric spaces*, Annals of Mathematics, (2), vol. 36 (1935), pp. 719–723.
8. G. O. Thorin, *An extension of a convexity theorem due to M. Riesz*, Fysiografiska Sällskapets i Lund, Förhandlingar (Proceedings of the Royal Physiographic Society at Lund), vol. 8 (1938), no. 14.

CAMBRIDGE, ENGLAND

NOTE ON DEGREE OF TRIGONOMETRIC AND POLYNOMIAL APPROXIMATION TO AN ANALYTIC FUNCTION, IN THE SENSE OF LEAST pTH POWERS

J. L. WALSH AND W. E. SEWELL

1. **Introduction.** In a recent note[1] the present writers studied the relation between the continuity properties of a function and the degree of approximation in the sense of Tchebycheff by trigonometric and other polynomials; this approximation in the z-plane is considered either on the unit circle by polynomials in z and $1/z$, or on the segment $-1 \leqq z \leqq 1$ by polynomials in z, or (for functions of period 2π) on the infinite interval $-\infty < z < \infty$ by trigonometric polynomials. In the respective cases, the functions approximated are analytic in an annulus $\rho > |z| > 1/\rho < 1$, in an ellipse whose foci are $+1$ and -1, or in a horizontal strip containing the axis of reals in its interior. It is the purpose of the present note to establish the analogous results when approximation is measured by the integral of the pth power of the error, as in the sense of least pth powers.

The method we employ makes essential use of the specific results concerning Tchebycheff approximation as developed in our previous note, together with certain general methods already developed elsewhere.[2]

2. **Approximation on the unit circle.** Our main result is as follows:

THEOREM 1. *Let the weight function $w(\theta)$ be positive and continuous for all values of θ, and of period 2π. Let the function $f(\theta)$ (not necessarily real) be periodic with period 2π, and suppose the numbers a_{nj} and b_{nj} (not necessarily real) are given so that*

$$s_n(\theta) = \frac{a_{n0}}{2} + \sum_{j=1}^{n} (a_{nj} \cos j\theta + b_{nj} \sin j\theta),$$

with the relation, for $n = 1, 2, \cdots$,

$$\int_{-\pi}^{\pi} w(\theta) \, | f(\theta) - s_n(\theta) |^p d\theta \leqq \frac{M}{\rho^{np} n^{(k+\alpha+1)p}}, \tag{1}$$

$$0 < \alpha \leqq 1, \; p > 0, \; \rho > 1,$$

[1] This Bulletin, vol. 44 (1938), pp. 865–873. We shall refer to this note as WS.

[2] Walsh, *Interpolation and Approximation by Rational Functions in the Complex Domain*, American Mathematical Society Colloquium Publications, vol. 20, New York, 1935.

where k is a nonnegative integer and M is a constant independent of n. Then the function

$$F(z) = \lim_{n\to\infty}\left[c_{n0} + \sum_{j=1}^{n}(c_{n,-j}z^{-j} + c_{nj}z^{j})\right], \tag{2}$$

$$2c_{nj} = a_{nj} - ib_{nj}, \qquad 2c_{n,-j} = a_{nj} + ib_{nj},$$

coincides with $f(\theta)$ almost everywhere (everywhere if $f(\theta)$ is continuous) on the circle $|z| = 1$, with $z = \cos\theta + i\sin\theta$, and $F(z)$ is analytic in the annulus $\rho > |z| > 1/\rho$ and continuous in the corresponding closed region. For z_1 and z_2 on $|z| = \rho$ or $1/\rho$ we have[3]

$$|F^{(k)}(z_1) - F^{(k)}(z_2)| \leqq L|z_1 - z_2|^{\alpha}\cdot|\log|z_1 - z_2||^{\beta}, \tag{3}$$

where $\beta = 0$ if $\alpha < 1$ and $\beta = 1$ if $\alpha = 1$, and where L is a constant independent of z_1 and z_2.

As a matter of convenience in the proof of Theorem 1 we establish a preliminary proposition:

LEMMA. *Let C be a rectifiable Jordan curve in the z-plane containing $z = 0$ in its interior; let $P(z, 1/z)$ be a polynomial in z and $1/z$ of degree n:*

$$P(z, 1/z) \equiv P_1(z) + P_2(z),$$

$$P_1(z) = a_0 + a_1z + \cdots + a_nz^n,$$

$$P_2(z) = a_{-1}z^{-1} + a_{-2}z^{-2} + \cdots + a_{-n}z^{-n},$$

and suppose

$$\int_C |P(z, 1/z)|^p|dz| \leqq N^p, \qquad p > 0. \tag{4}$$

Denote by $w = f_1(z)$ a function which maps the exterior of C onto $|w| > 1$ so that $f_1(\infty) = \infty$, and denote by $w = f_2(z)$ a function which maps the interior of C onto $|w| < 1$ so that $f_2(0) = 0$; denote by C_R the closed annular region bounded by the two Jordan curves $|f_1(z)| = R > 1$, $|f_2(z)| = 1/R$.

Then there exists a constant N' depending on C, p, and R but not on $P(z, 1/z)$ nor on n, such that (4) *implies*

$$|P(z, 1/z)| \leqq NN'R^n, \quad z \text{ in } C_R.$$

By a known method of proof (Walsh, op. cit., p. 92) we have for suitably chosen N'

[3] The notation $F^{(k)}(z)$ indicates the kth derivative of $F(z)$ if $k > 0$ and the function $F(z)$ itself for $k = 0$. Here and below such derivatives on the boundaries of regions of analyticity are considered in the one-dimensional sense.

$$|P(z, 1/z)| \leqq NN'R^n, \quad z \text{ on curve } |f_1(z)| = R;$$
$$|P(z, 1/z)| \leqq NN'R^n, \quad z \text{ on curve } |f_2(z)| = 1/R.$$

The lemma follows at once.[4]

We proceed now with the proof of Theorem 1. From (1) by the well known general inequalities

$$(5) \qquad \begin{aligned} |\chi_1 + \chi_2|^p &\leqq 2^{p-1}|\chi_1|^p + 2^{p-1}|\chi_2|^p, && p > 1, \\ |\chi_1 + \chi_2|^p &\leqq |\chi_1|^p + |\chi_2|^p, && 0 < p \leqq 1, \end{aligned}$$

and by virtue of the boundedness of $1/w(\theta)$, we have

$$(6) \qquad \int_{-\pi}^{\pi} |s_{n+1}(\theta) - s_n(\theta)|^p d\theta \leqq \frac{M_1}{\rho^{np} n^{(k+\alpha+1)p}}.$$

On the circle $|z| = 1$ we have $\cos j\theta = (z^j + z^{-j})/2$, $\sin j\theta = (z^j - z^{-j})/2i$, so the first member of (6) can be written

$$\int_{|z|=1} |P_{n+1}(z, 1/z) - P_n(z, 1/z)|^p |dz|,$$

where $P_n(z, 1/z)$ is a polynomial of degree n in z and $1/z$, equal to $s_n(\theta)$ on $|z| = 1$. From inequality (6) we now deduce through the lemma for $\rho \geqq |z| \geqq 1/\rho$,

$$(7) \qquad |P_{n+1}(z, 1/z) - P_n(z, 1/z)| \leqq M_2/n^{k+\alpha+1},$$

where M_2 is independent of n and z. We define $F(z)$ for $\rho \geqq |z| \geqq 1/\rho$ by means of the equation

$$(8) \qquad \begin{aligned} F(z) \equiv P_1(z, 1/z) &+ [P_2(z, 1/z) - P_1(z, 1/z)] \\ &+ [P_3(z, 1/z) - P_2(z, 1/z)] + \cdots, \end{aligned}$$

whence from (7) we have for $\rho \geqq |z| \geqq 1/\rho$

$$(9) \qquad \begin{aligned} |F(z) - P_n(z, 1/z)| &\leqq |P_{n+1}(z, 1/z) - P_n(z, 1/z)| \\ &\quad + |P_{n+2}(z, 1/z) - P_{n+1}(z, 1/z)| + \cdots \\ &\leqq M_2 \sum_{j=n}^{\infty} \frac{1}{j^{k+\alpha+1}} \leqq \frac{M_3}{n^{k+\alpha}}, \end{aligned}$$

where M_3 is independent of n and z.

The uniformity of the convergence in (8) is included in the inequalities (9), so it follows that $F(z)$ is analytic for $\rho > |z| > 1/\rho$, continuous

[4] So far as the writers are aware, this explicit lemma has not been previously formulated in print, but indications of it are given by Walsh, American Journal of Mathematics, vol. 54 (1932), pp. 559–570.

for $\rho \geqq |z| \geqq 1/\rho$; and since $P_n(z, 1/z)$ on $|z| = \rho$ and $|z| = 1/\rho$ is a trigonometric polynomial of order n, it follows from a theorem due to de la Vallée Poussin[5] that $F(z)$ satisfies an inequality *with respect to* θ of form (3), $z = \rho e^{i\theta}$ or $e^{i\theta}/\rho$, from which it follows that inequality (3) itself is fulfilled on $|z| = \rho$ and $|z| = 1/\rho$.

The boundedness of $1/w(z)$ implies from (1)

$$\int_{-\pi}^{\pi} |f(\theta) - s_n(\theta)|^p d\theta \leqq \frac{M_4}{\rho^{np} n^{(k+\alpha+1)p}}.$$

Inequality (9) implies, from the equality of $P_n(z, 1/z)$ and $s_n(\theta)$ on $|z| = 1$,

$$\int_{-\pi}^{\pi} |F(z) - s_n(\theta)|^p d\theta \leqq \frac{M_5}{n^{(k+\alpha)p}},$$

so we now deduce by use of (5) and by allowing n to become infinite

$$\int_{-\pi}^{\pi} |F(z) - f(\theta)|^p d\theta = 0;$$

hence $F(z)$ and $f(\theta)$ are equal almost everywhere on $|z| = 1$, and Theorem 1 is established.

3. **Approximation on the segment** $-1 \leqq z \leqq 1$. We shall indicate rapidly the proof of the next theorem.

THEOREM 2. *Let $f(z)$ be defined on the segment $-1 \leqq z \leqq 1$, and for $n = 1, 2, \cdots$ let a polynomial $P_n(z)$ in z of degree n exist such that*

$$\int_{-1}^{1} w(z) |f(z) - P_n(z)|^p dz \leqq \frac{M}{\rho^{np} n^{(k+\alpha+1)p}}, \tag{10}$$

$$0 < \alpha \leqq 1, \ p > 0, \ \rho > 1,$$

where k is a nonnegative integer, and where $w(z)(1-z^2)^{1/2}$ is positive and continuous on $-1 \leqq z \leqq 1$. Then the function $f(z)$ is equal to $F(z)$ almost everywhere on the segment $-1 \leqq z \leqq 1$, where

$$F(z) = \lim_{n \to \infty} P_n(z). \tag{11}$$

Furthermore $F(z)$ is analytic throughout the interior of the ellipse γ whose foci are $+1$ and -1 and the sum of whose semi-axes is ρ; also $F(z)$ is continuous in the corresponding closed region, and satisfies an inequality of type (3) on γ.

[5] *Leçons sur l'Approximation*, Paris, 1919, chap. 4.

We map the z-plane onto the w-plane by the transformation $z=(w+w^{-1})/2$, which carries the segment $-1\leqq z\leqq 1$ counted twice into the unit circle $|w|=1$, and carries the ellipse γ counted twice into the two circles $|w|=\rho$, $|w|=1/\rho$. The first member of (10) can be written

$$\frac{1}{2}\int_{|w|=1} w(z)\,|f(z)-P_n(z)|^p|(1-z^2)^{1/2}|\,|dw|,$$

so Theorem 1 applies (cf. WS for further details of the transformation and reasoning). A condition of form (3) on $|w|=\rho$ and $|w|=1/\rho$ with respect to w implies condition (3) as stated in Theorem 2, so the conclusion follows.

Another method of studying the situation of Theorem 2 is to avoid the transformation onto the w-plane, but to use the fact that for a polynomial $P(z)$ of degree n the inequality $\int_{-1}^{1}|P(z)|^p dz\leqq N^p$, $p>0$, implies $|P(z)|\leqq NN'R^n$ on and within[6] γ. Here it is natural to assume that $w(z)$ itself is positive and continuous rather than $w(z)(1-z^2)^{1/2}$. We deduce as in the proof of Theorem 1 an inequality analogous to (9) giving the degree of convergence of the sequence $P_n(z)$ to $F(z)$ on γ; this degree of convergence then implies[7] inequality (3) on γ, as we wish to prove. Thus we have proved the following theorem:

THEOREM 2a. *Theorem 2 remains true if the requirement that $w(z)(1-z^2)^{1/2}$ is positive and continuous on $-1\leqq z\leqq 1$ is replaced by the requirement that $w(z)$ be positive and continuous on $-1\leqq z\leqq 1$.*

4. **Approximation to a periodic function on the axis of reals.** A different transformation of the plane will now yield a new result:

THEOREM 3. *Let the function $f(z)$ be periodic with period 2π, and for $n=1, 2, \cdots$ let there exist a trigonometric polynomial $t_n(z)$ of order n such that we have*

$$\int_{-\pi}^{\pi} w(z)\,|f(z)-t_n(z)|^p dz\leqq \frac{M}{\rho^{np}n^{(k+\alpha+1)p}},\quad 0<\alpha\leqq 1,\ p>0,\ \rho>1,$$

where the weight function $w(z)$ is a positive, continuous function of period 2π, and k is a nonnegative integer. Then if we define $F(z)=\lim_{n\to\infty} t_n(z)$, where $t_n(z)$ is still expressed as a trigonometric polynomial, the two functions $f(z)$ and $F(z)$ are equal almost everywhere

[6] Walsh, op. cit., Lemma, p. 92.

[7] Walsh and Sewell, this Bulletin, vol. 43 (1937), pp. 557–563. This method of proof of Theorem 2a is used widely in a number of similar situations in a forthcoming paper by the present writers, on "Problem β."

on the axis of reals. Moreover if $z=x+iy$, the function $F(z)$ is analytic for $|y|<\log \rho$, continuous in the corresponding closed region, and on the lines $y=\pm\log \rho$ satisfies an inequality of type (3).

Theorem 3 is a direct consequence of Theorem 1, by virtue of the transformation $w=e^{iz}$, in the notation of Theorem 3 (compare WS, proof of Theorem 5).

5. **Convergence properties of sequences.** The methods of proof of Theorems 1, 2, and 3 apply also to the study of the degree of convergence of the sequences involved on various point sets not yet mentioned. Thus under the hypothesis of Theorem 1, the method already used shows that in the annulus $\rho_1 \geqq |z| \geqq 1/\rho_1$, $\rho_1<\rho$, we have instead of (7)

$$|P_{n+1}(z, 1/z) - P_n(z, 1/z)| \leqq \frac{M_2\rho_1^n}{\rho^n n^{k+\alpha+1}}.$$

Then for $\rho_1 \geqq |z| \geqq 1/\rho_1$ we may write

$$\begin{aligned}|F(z) - P_n(z, 1/z)| &\leqq |P_{n+1}(z, 1/z) - P_n(z, 1/z)| \\ &\quad + |P_{n+2}(z, 1/z) - P_{n+1}(z, 1/z)| + \cdots \\ &\leqq M_2\sum_{j=n}^{\infty} \frac{\rho_1^j}{\rho^j j^{k+\alpha+1}} \leqq \frac{M_2}{n^{k+\alpha+1}} \sum_{j=n}^{\infty}\left(\frac{\rho_1}{\rho}\right)^j \leqq \frac{M'\rho_1^n}{\rho^n n^{k+\alpha+1}}.\end{aligned}$$

That is to say, we have established a further result:

COROLLARY TO THEOREM 1. *Under the hypothesis of Theorem* 1 *we have in the annulus $\rho_1 \geqq |z| \geqq 1/\rho_1$, $\rho_1<\rho$,*

$$|F(z) - P_n(z, 1/z)| \leqq \frac{M'\rho_1^n}{\rho^n n^{k+\alpha+1}},$$

where M' is a suitably chosen constant independent of n and z.

The corollary just established implies corresponding statements connected with Theorems 2 and 3.

COROLLARY TO THEOREM 2 (OR THEOREM 2a). *Under the hypothesis of Theorem* 2 (*or Theorem* 2a) *we have for z on and within the ellipse whose foci are $+1$ and -1 and whose sum of semi-axes is $\rho_1<\rho$*

$$|F(z) - P_n(z)| \leqq \frac{M'\rho_1^n}{\rho^n n^{k+\alpha+1}},$$

where M' is a suitably chosen constant independent of n and z.

In the proof of the Corollary to Theorem 2a we employ the method of proof of Theorem 2a rather than the Corollary to Theorem 1 itself.

COROLLARY TO THEOREM 3. *Under the hypothesis of Theorem 3 we have for z in the region $|y| < \log \rho_1 < \log \rho$*

$$|F(z) - t_n(z)| \leqq \frac{M'\rho_1^n}{\rho^n n^{k+\alpha+1}},$$

where M' is a suitably chosen constant independent of n and z.

It may be noted that the inequalities representing the conclusions of these corollaries, with the obvious understanding that these functions $P_n(z, 1/z)$, $P_n(z)$, $t_n(z)$ are polynomials of the kind considered of respective degree n, themselves imply the conclusions of Theorems 1, 2 (or 2a), and 3 (compare WS, Theorems 1, 3, and 5).

6. **Reciprocal theorems.** In the direction of a converse of Theorem 1 we indicate the proof of the following result:

THEOREM 4. *Let the function $F(z)$ be analytic for $\rho > |z| > 1/\rho < 1$, continuous in the corresponding closed region, and let $F^{(k)}(z)$, k a nonnegative integer, satisfy a Lipschitz condition[8] of order α on $|z| = \rho$ and $|z| = 1/\rho$. Let $p > 0$ be given. Let $w(\theta)$ be a nonnegative function of θ, periodic with period 2π, and Lebesgue-integrable for $-\pi \leqq \theta \leqq \pi$. Then there exists a sequence of polynomials $P_n(z, 1/z)$ in z and $1/z$ of degrees $n = 1, 2, \cdots$ so that we have*

$$\int_{-\pi}^{\pi} w(\theta) |F(z) - P_n(z, 1/z)|^p d\theta \leqq \frac{M}{\rho^{np} n^{(k+\alpha)p}},$$

where M is a suitably chosen constant independent of n and z.

Theorem 4 is an immediate consequence of the corresponding theorem for Tchebycheff approximation (WS, Theorem 2). We add the remark that the conclusion of Theorem 4, holding for some sequence $P_n(z, 1/z)$, holds *a fortiori* for the sequence $P_n(z, 1/z)$ of best approximation in the sense of least pth powers. The Corollary to Theorem 1 applies to the latter.

Corresponding results in the directions of converses of Theorems 2, 2a, and 3 can be proved with no less ease.

It will be noticed that there is a discrepancy of unity between the exponents of n in Theorems 1 and 4. This discrepancy is not acci-

[8] That is to say, let (3) be satisfied with $\beta = 0$.

dental, but is inherent in the nature of the problem, as is shown by examples that the authors plan to publish elsewhere.

7. **Generality of weight functions.** The mere casual comparison of Theorems 2 and 2a suggests that the particular restrictions we have placed on the weight functions are artificial rather than essential to the problem. The most general weight function that can be employed seems to be difficult to determine, both here and in the case of polynomial approximation purely in the real domain. Nevertheless we shall prove the following additional proposition:

THEOREM 5. *Theorems* 1, 2, 2a, *and* 3 *persist if the requirement that the respective weight functions* $w(\theta)$, $w(z)(1-z^2)^{1/2}$, $w(z)$, $w(z)$ *be positive and continuous is replaced by the requirement that the weight functions be nonnegative, Lebesgue-integrable, and that some negative power of the corresponding weight functions be Lebesgue-integrable.*[9]

The situation of Theorem 1 is typical. Under the new assumption of Theorem 5 we assume $[w(\theta)]^{-\beta}$ to be integrable in the interval $-\pi \leqq \theta \leqq \pi$, with $\beta > 0$. The Hölder inequality

$$\left| \int F^{\delta} G^{1-\delta} \right| \leqq \left(\int |F| \right)^{\delta} \left(\int |G| \right)^{1-\delta}, \qquad 0 < \delta < 1,$$

with $\delta = 1/(1+\beta)$, gives

$$\int_{-\pi}^{\pi} |f(\theta) - s_n(\theta)|^{p(1-\delta)} d\theta \leqq \left(\int_{-\pi}^{\pi} \frac{d\theta}{[w(\theta)]^{(1-\delta)/\delta}} \right)^{\delta} \left(\int_{-\pi}^{\pi} w(\theta) |f(\theta) - s_n(\theta)|^p d\theta \right)^{1-\delta},$$

and by virtue of (1) we have

$$\int_{-\pi}^{\pi} |f(\theta) - s_n(\theta)|^{p(1-\delta)} d\theta \leqq \frac{M'}{\rho^{np(1-\delta)} n^{(k+\alpha+1)p(1-\delta)}}.$$

This last inequality leads to an inequality similar to (6) and thence as before to the conclusion of Theorem 1.

HARVARD UNIVERSITY AND
GEORGIA SCHOOL OF TECHNOLOGY

[9] This condition on the weight functions has been previously used by Dunham Jackson in the study of approximation in the real domain, and by Walsh (op. cit., pp. 104–105) in the complex domain.

A RELATION BETWEEN NON-ALTERNATING AND INTERIOR TRANSFORMATIONS[1]

G. T. WHYBURN

Recently in proving certain existence theorems for non-alternating and for interior transformations of a continuum onto a simple arc[2] it was observed that when a transformation of one of these types was set up, usually it satisfied in large measure conditions which brought it also under transformations of the other type. This suggests the existence of common ground shared by these sorts of transformations, and it is the object of this paper to exhibit the nature of such. Our principal conclusion is to the effect that, under certain auxiliary conditions, any interior transformation $f(M)=D$ of a compact continuum M onto a dendrite D can be factored into the form $f=f_2f_1$ where $f_1(M)=M'$ merely shrinks sets of type $f^{-1}(p)$ (p an end point of D) into single points and is topological elsewhere and $f_2(M')=D$ is non-alternating and interior. However, we first prove two theorems on the relation between the non-alternating property of a transformation of a continuum into a dendrite and the connectedness of the inverse sets for the end points of the dendrite.

THEOREM 1. *If M is a compact metric continuum and $f(M)=N$ is non-alternating, then for any end point[3] p of N, $f^{-1}(p)$ is connected.*

PROOF. Suppose $f^{-1}(p)$ is not connected. Then there exist points a and b of $f^{-1}(p)$, a closed set $F \subset M-f^{-1}(p)$ and a separation $M-F=M_a+M_b$ where $a \in M_a$, $b \in M_b$. Since $f(F)\cdot p=0$, $f(F)$ is compact, and p is an end point, there exists a neighborhood U of p in N with $\overline{U}\cdot f(F)=0$ and such that the boundary of U is a single point q. Since q separates p and $f(F)$ in N, it follows by a result of the author's[4] that there exists a separation $M-f^{-1}(q)=M_F+M_p$, where $F \subset M_F$, $f^{-1}(p) \subset M_p$. But this gives at once the separation $M-f^{-1}(q)=M_a\cdot M_p+(M_F+M_b\cdot M_p)$ where $M_a\cdot M_p \supset a$, $(M_F+M_b\cdot M_p) \supset b$, contrary to the hypothesis that f is non-alternating.

THEOREM 2. *If $f(M)=D$ is an interior transformation of a compact*

[1] Presented to the Society, September 8, 1939.

[2] See my papers *The existence of certain transformations*, Duke Mathematical Journal, vol. 5 (1939), pp. 647–655; and *Non-alternating interior retracting transformations*, Annals of Mathematics, (2), vol. 40 (1939), pp. 914–921.

[3] That is, a point of Menger-Urysohn order 1 of N.

[4] See my paper in the American Journal of Mathematics, vol. 56 (1934), pp. 294–302, (1.4).

continuum M onto a dendrite D such that for each end point p of D, $f^{-1}(p)$ is connected, then f is non-alternating.

PROOF. Let $x, y \in D$ and let R be the component of $D-x$ containing y. Then R contains at least one end point p of D. Since $f^{-1}(p)$ is connected, there exists a single component Q of $M-f^{-1}(x)$ containing $f^{-1}(p)$. Since,[5] by the interiority of f, any quasi-component of $f^{-1}(R)$ maps onto all of R and hence intersects $f^{-1}(p)$, we have $Q=f^{-1}(R)$, whence $Q \supset f^{-1}(y)$ and thus f is non-alternating.

COROLLARY. *In order that an interior transformation $f(M)=D$ of a compact continuum M onto a dendrite D be non-alternating it is necessary and sufficient that for each end point p of D, $f^{-1}(p)$ be connected.*

THEOREM 3. *Let M be a compact continuum, let D be a dendrite and let H be the set of all end points of D. Any interior transformation $f(M)=D$ such that the collection $[f^{-1}(p)]$, $p \in H$, is semi-closed*[6] *can be factored into the form $f=f_2f_1$, where f_1 is equivalent to f on $f^{-1}(H)$ and topological on $M-f^{-1}(H)$ and f_2 is interior and non-alternating.*

PROOF. Let us decompose M into the sets $f^{-1}(p)$, $p \in H$, and individual points of $M-f^{-1}(H)$. Since the collection $[f^{-1}(p)]$, $p \in H$, is semi-closed, it follows by a result of the author's[6] that this decomposition is upper semi-continuous.[7] Let M' be the hyperspace of this decomposition and $f_1(M)=M'$ the associated continuous transformation.

Now for each $x' \in M'$, let us define $f_2(x')=ff_1^{-1}(x')$. Then clearly $f(x)\equiv f_2f_1(x)$ for $x \in M$. The continuity of f_2 results at once from that of f_1 and f. Furthermore, since f is interior it follows[8] that f_2 is interior. Finally, since for any end point p of D we have that $f_2^{-1}(p)=f_1f^{-1}(p)$ is a single point of M', it follows by Theorem 2 that f_2 is non-alternating. This completes the proof of our theorem.

Since any finite collection of disjoint closed sets is semi-closed, we have the following corollary:

COROLLARY. *Any interior transformation of a compact continuum into a simple arc (or into a dendrite with only a finite number of end points) can be factored into the form described in Theorem 3.*

UNIVERSITY OF VIRGINIA

[5] See my paper, Duke Mathematical Journal, vol. 3 (1937), pp. 370–381, (1.4).

[6] That is, any convergent sequence of the sets $[f^{-1}(p)]$ whose limit set intersects $M-f^{-1}(H)$ converges to a single point. See my paper in the Duke Mathematical Journal, vol. 2 (1937), pp. 685–690.

[7] See R. L. Moore, *Foundations of Point Set Theory*, American Mathematical Society Colloquium Publications, vol. 13, 1932, chap. 5.

[8] See result (1.6) of the paper cited in footnote 5.

CONCERNING SIMILARITY TRANSFORMATIONS OF LINEARLY ORDERED SETS[1]

BEN DUSHNIK AND E. W. MILLER

1. **Introduction.** As is well known, two linearly ordered sets A and B are said to be similar if there exists a 1-1 correspondence between their elements which preserves order. A function f which defines such a 1-1 correspondence may be called a similarity transformation on A to B. In this note we consider two problems concerning similarity transformations which do not seem to have received attention heretofore. The first problem is the following:

(A) *Is it true that every infinite ordered set is similar to a proper subset of itself?*[2]

Before stating the second problem we recall a classical theorem concerning well-ordered sets.[3]

THEOREM. *If the set A is well-ordered, and if f is any similarity transformation on A to a subset of A, then $f(a) \geqq a$ for every a in A.*

It is natural to inquire whether this theorem characterizes well-ordered sets—and this is our second problem; more explicitly:

(B) *Let A be a linearly ordered set such that if f is any similarity transformation on A to a subset of A then $f(a) \geqq a$ for every a in A. Is it true that any such set A is well-ordered?*

We will demonstrate that if the set A is denumerable, then the answer to both questions is in the affirmative. An example will then be constructed to show that these conclusions need not hold if the set A is nondenumerable.

2. **The denumerable case.** We obtain first the following result:

THEOREM 1. *Every denumerably infinite linearly ordered set A contains a proper subset A' to which it is similar.*

PROOF. For any two elements a and b of A, we will say that a and b are *congruent* if either $a=b$ or if there is only a finite number of ele-

[1] Presented to the Society, September 8, 1939.

[2] This question is a natural one, in view of the familiar definition of an infinite set as one which is *equivalent* to a proper subset of itself.

[3] For theorems mentioned in this paper one may refer to Hausdorff's *Grundzüge der Mengenlehre*, 1st edition, 1914, or to Sierpiński's *Leçons sur les Nombres Transfinis*, 1928.

ments in A which lie between a and b; we will indicate this by writing $a \equiv b$. The set of all elements of A congruent to a given element a will be designated by $C(a)$ and will be called the *congruence set* corresponding to a. It is obvious that if $b \in C(a)$, then $C(b) = C(a)$, and hence that any two different congruence sets have no elements in common. We now distinguish two cases.

Case 1. There is an a in A for which $C(a)$ is infinite.

Since every element of $C(a)$ has either an immediate predecessor or an immediate successor in $C(a)$, it is clear that the order-type of $C(a)$ is either ω or ω^* or $\omega^* + \omega$. Suppose, for example, that the order-type of $C(a)$ is ω. Since every element of A which is not in $C(a)$ either precedes or succeeds *all* the elements of $C(a)$, we will have $A = A_1 + A_2 + A_3$, where $A_2 = C(a)$, A_1 is the set of all elements of A which precede $C(a)$, and A_3 is the set of all the elements of A which follow $C(a)$. We now define a function f on A to a subset of A as follows:

(a) If $\alpha \in A_1$, or $\alpha \in A_3$, then $f(\alpha) = \alpha$.

(b) If $\alpha \in A_2$, then $f(\alpha)$ is the successor of α in $C(a)$.

This function is clearly a similarity transformation; moreover, the set A' into which A is transformed by f does not contain the first element of $C(a)$, and A' is thus a proper subset of A. The case where the order-type of $C(a)$ is ω^* or $\omega^* + \omega$ may be handled in an analogous fashion.

Case 2. All the different congruence sets in A are finite sets.

Since A is denumerable, the set of all congruence sets in A is also denumerable. Let $\overline{A}$ be the ordered subset of A which consists of all first elements of congruence sets in A, and let α and β, $(\alpha < \beta)$, be any two elements of $\overline{A}$. There must exist a γ in $\overline{A}$ for which $\alpha < \gamma < \beta$ (for, if not, then clearly $\alpha \equiv \beta$ in A, which would contradict the fact that α and β belong to two different congruence sets). Thus $\overline{A}$ is a denumerable dense set, and any open interval of it, for example, the set (α, β) of all elements of $\overline{A}$ between α and β, would also be a denumerable dense set. It is well known[4] that such a set contains a subset similar to any given denumerable ordered set. Let A' be a subset of (α, β) which is similar to A. It is clear that A' is a proper subset of A, and Theorem 1 is thus completely demonstrated.

THEOREM 2. *If the denumerably infinite ordered set A is such that any similarity transformation f on A to a subset of A has the property $f(a) \geqq a$ for every a in A, then A is a well-ordered set.*

PROOF. We notice at once that A cannot contain a subset $\overline{A}$ which

[4] See, for example, Sierpiński, loc. cit., pp. 147–148.

is dense. For, if $\overline{A}$ is such a subset, let $\overline{\overline{A}}$ be the lower segment of $\overline{A}$ which consists of all the elements of $\overline{A}$ which precede some element b in $\overline{A}$. Then $\overline{\overline{A}}$ would be a denumerable dense set, and, as was pointed out in the proof of Theorem 1, $\overline{\overline{A}}$ would contain a subset A' similar to A, that is, there would exist a similarity transformation f on A to A'. Since $f(b) \in A'$, and all elements of A' precede b, we would obtain $f(b) < b$; but this is impossible, by our hypothesis.

We now redefine "congruent" elements of A as follows: $a \equiv b$ if and only if (1) $a = b$ or (2) the closed interval of A which consists of a and b and all the elements of A between a and b is, as an ordered subset of A, a well-ordered set. In the present sense, the congruence sets have the following properties:

(1) Two different congruence sets have no elements in common.

(2) Any upper segment, or any open or closed interval of a congruence set, is a well-ordered set.

(3) If $a \notin C(b)$, then a either precedes all, or succeeds all of the elements of $C(b)$.

It may be emphasized, however, that a lower segment of a congruence set is well-ordered only if the congruence set has a first element. We again distinguish two cases.

Case 1. Every congruence set in A has a first element, that is, every congruence set is well-ordered.

In this case, we will show that there can be only one congruence set altogether; the set A will thus be identical with this unique congruence set, and will therefore be well-ordered. To see that we cannot have more than one congruence set here, suppose the contrary, and let α and β, $(\alpha < \beta)$, be the first elements of any two different congruence sets, which sets may be designated by $C(\alpha)$ and $C(\beta)$. If there were in A no element c which separates $C(\alpha)$ and $C(\beta)$, then the set $D = C(\alpha) + C(\beta)$, considered as an ordered subset of A, would be an interval of A which is well-ordered. This would mean that $\alpha \equiv \beta$; but this cannot be, since α and β belong to different congruence sets. Hence, there must exist a congruence set $C(\gamma)$, whose first element is γ, such that $C(\alpha) < C(\gamma) < C(\beta)$, and we would thus have $\alpha < \gamma < \beta$. This means that the set A' of all the first elements of congruence sets in A will be a dense set—contrary to the fact that A cannot contain such a subset.

Case 2. There is a congruence set C without a first element.

We will show that this supposition leads to a contradiction, so that this case cannot actually arise. The proof of our theorem will then be complete.

We can find in C a sequence of elements $c_1 > c_2 > c_3 > \cdots > c_n > \cdots$

such that for every d in C there is a natural number n for which $c_n \leqq d$. The set of elements $[c_{i+1}, c_i)$ of C between c_i and c_{i+1}, inclusive of c_{i+1}, is a well-ordered set; let β_i be the order-type of this set, and consider the sequence of ordinals $\beta_1, \beta_2, \cdots, \beta_n, \cdots$ thus defined. There can be only a finite number of indices k such that for all $n > k$, $\beta_n < \beta_k$ (otherwise one would obtain an infinite sequence of decreasing ordinals, which is impossible). Hence, there will exist an index r such that for every $p \geqq r$ there will be infinitely many indices $q > p$ for which $\beta_p \leqq \beta_q$. In other words, for $i = r, r+1, r+2, \cdots$ the set $[c_{i+1}, c_i)$ will be similar to a subset of some $[c_{q_i+1}, c_{q_i})$ where $q_i > i$, and where moreover, $q_r < q_{r+1} < q_{r+2} < \cdots$. We now express A as the sum of three sets A_1, A_2 and A_3, where A_2 consists of all the elements of C preceding c_r, A_1 consists of all elements of A preceding every element of A_2, and A_3 consists of all the elements of A following every element of A_2. Define now a function f on A to a subset of A as follows: If $\alpha \in A_1$ or $\alpha \in A_3$, $f(\alpha) = \alpha$. Suppose that $\alpha \in A_2$. There will be an $i \geqq r$ for which $\alpha \in [c_{i+1}, c_i)$, and since $[c_{i+1}, c_i)$ is similar to a subset of $[c_{q_i+1}, c_{q_i})$, there will be a γ_α in this last set which corresponds to α under such a similarity. We define, for $\alpha \in A_2$, $f(\alpha) = \gamma_\alpha$.

This function is clearly a similarity transformation on A to a subset of itself, and for any α in A_2 we have $f(\alpha) = \gamma_\alpha < \alpha$; but this is impossible, by our hypothesis.

It may be observed that up to this point it has not been necessary to make use of Zermelo's axiom of choice.

3. **The nondenumerable case.** We now prove the following result:

THEOREM 3. *The linear continuum C contains a set E, of power $\mathfrak{c}$, which is not similar to any proper subset of itself.*

PROOF. A similarity transformation of C into a subset of itself is a monotonic increasing function of a real variable, and vice versa. Hence, there are exactly $\mathfrak{c}$ such transformations. Let $\Omega_\mathfrak{c}$ denote the first ordinal to correspond to the cardinal number of the continuum, and arrange all of these transformations, with the exception of the identity, in a well-ordered series of type $\Omega_\mathfrak{c}$: $T_1, T_2, \cdots, T_\alpha, \cdots$, $(\alpha < \Omega_\mathfrak{c})$.

Now, the fixed points under a monotonic transformation which is not the identity cannot be everywhere dense in C. Hence, if T is such a transformation, there are $\mathfrak{c}$ points p such that $T(p) \neq p$. We may accordingly choose a point p_1 such that $T_1(p_1) = q_1 \neq p_1$. We now assume that p_β and q_β have been defined for all $\beta < \alpha < \Omega_\mathfrak{c}$, and choose distinct points p_α and q_α in $C - \sum_{\beta<\alpha}(p_\beta + q_\beta)$ such that $T_\alpha(p_\alpha) = q_\alpha$.

This is possible for every $\alpha < \Omega_c$, since the power of $\sum_{\beta<\alpha}(p_\beta + q_\beta)$ is less than $\mathfrak{c}$.

We notice first that if C' is any interval on C, then there exists a monotonic increasing function which is the identity on $C - C'$ and is not the identity on C'. Therefore there is an α such that $p_\alpha \in C'$. Hence the set $E = \sum_{\alpha<\Omega_c} p_\alpha$ is dense on C. It now follows that there is no similarity transformation of E into a proper subset of itself. In fact, we can show that the only similarity transformation of E into a subset of itself is the identity. For assume that τ is a similarity transformation, other than the identity, of E into a subset of itself. Since τ is not the identity and since E is dense in C, there exists an α such that T_α agrees with τ on E. But this is impossible, since $T_\alpha(p_\alpha) = q_\alpha$, and q_α is not a point of E.

THEOREM 4. *The linear continuum contains a set E which is not well-ordered, but which has the property that if f is any similarity transformation of E into a subset of itself, then $f(e) \geqq e$ for every e in E.*

PROOF. The set E constructed in the proof of Theorem 3 satisfies these requirements.

UNIVERSITY OF MICHIGAN

ON PERFECT SUMMABILITY OF DOUBLE SEQUENCES

J. D. HILL

1. Introduction. The purpose of the present note is to point out that the results of Banach on perfect summability of simple sequences[1] may be extended to certain cases of double sequence summability. The main result obtained is embodied in the theorem of §3.

It will be convenient to begin with the following definitions and notations. We denote by c the class of all real double sequences $x \equiv \{\xi_{kl}\}$ for which the *principal limit* $\lim_{k,l} \xi_{kl} \equiv \xi$ exists finite; by rc, the subclass of c for which the *row limits* $\lim_l \xi_{kl} \equiv \xi_k^r$, $(k=0, 1, 2, \cdots)$ and the *column limits* $\lim_k \xi_{kl} \equiv \xi_l^c$, $(l=0, 1, 2, \cdots)$, exist finite; and by $rcrn$, the subclass of rc throughout which $\xi_k^r = \xi_l^c = \xi = 0$, $(k, l=0, 1, 2, \cdots)$. With the conventional definitions of addition and multiplication by a constant the classes rc and $rcrn$ are linear spaces, and they become Banach spaces upon introduction of the norm $\|x\| \equiv \sup_{k,l} |\xi_{kl}|$.

Let $A \equiv (a_{ijkl})$ be a given infinite matrix of real numbers. We shall be concerned with transformations of the form

$$A_{ij}(x) \equiv \sum_{k,l} a_{ijkl}\xi_{kl}, \qquad i, j = 0, 1, 2, \cdots, \tag{1.1}$$

on the elements $x \equiv \{\xi_{kl}\}$ of rc. More precisely, if $A_{ij}(x)$, $(i, j = 0, 1, 2, \cdots)$, exists for every $x \in rc$ and if the corresponding sequence $\{A_{ij}(x)\}$ belongs to rc $[c]$, one says that rc is transformed into rc $[c]$ by the method of summability corresponding to the matrix A, or simply, by the method A. If $A(x) \equiv \lim_{i,j} A_{ij}(x) = \xi$ for every $x \in rc$, the method is called *regular*. A regular transformation of rc into itself is called *completely regular* if it is also *regular by rows and columns*, that is to say, if $A_i^r(x) \equiv \lim_j A_{ij}(x) = \xi_i^r$, $(i=0, 1, 2, \cdots)$, and $A_j^c(x)$ $\lim_i A_{ij}(x) = \xi_j^c$, $(j=0, 1, 2, \cdots)$, for every $x \in rc$.

For later reference we recall that the condition

$$\sup_{i,j} \sum_{k,l} |a_{ijkl}| \equiv M < \infty \tag{1.2}$$

is necessary[2] in order that A transform $rcrn$ into rc.

If the system of equations

$$\sum_{k,l} a_{ijkl}\xi_{kl} = \eta_{ij}, \qquad i, j = 0, 1, 2, \cdots, \tag{1.3}$$

[1] Banach, *Théorie des Opérations Linéaires*, pp. 90–95.

[2] See Hamilton, *Transformations of multiple sequences*, Duke Mathematical Journal, vol. 2 (1936), pp. 29–60; in particular, p. 42, #5.

has a unique solution $\{\xi_{kl}\}$ (belonging to rc or not) corresponding to each $\{\eta_{ij}\} \in rc$, the method A will be called *reversible.*[3]

REMARK. If A is reversible, it can be shown[4] that there exist linear functionals $f_{kl}(y)$ defined in rc, such that for each $y \equiv \{\eta_{ij}\} \in rc$ we have $\sum_{k,l} a_{ijkl} f_{kl}(y) = \eta_{ij}$, $(i, j = 0, 1, 2, \cdots)$. This fact is essential to the proof of Lemma 4 in §3.

A will be said to be of *type* M if the conditions[5]

$$\sum_{i,j} |\alpha_{ij}| < \infty, \qquad \sum_{i,j} \alpha_{ij} a_{ijkl} = 0, \qquad k, l = 0, 1, 2, \cdots, \tag{1.4}$$

always imply $\alpha_{ij} = 0$, $(i, j = 0, 1, 2, \cdots)$.

Any method A which is simultaneously reversible, completely regular, and of type M will be called *perfect.*[6]

Frequent use will be made of the following particular double sequences: the sequence X, all of whose terms are 1; the sequence X_k^r, $(k = 0, 1, 2, \cdots)$, all of whose terms are 0 except those in the kth row which are 1; the sequences X_l^c, $(l = 0, 1, 2, \cdots)$, all of whose terms are 0 except those in the lth column which are 1; and the sequences X_{kl}, $(k, l = 0, 1, 2, \cdots)$, all of whose terms are 0 except for the term common to the kth row and lth column, which is 1.

Corresponding to any given $x \in rc$ we set

$$\begin{aligned} x_i \equiv \xi X &+ \sum_{k=0}^{p_i} (\xi_k^r - \xi) X_k^r + \sum_{l=0}^{p_i} (\xi_l^c - \xi) X_l^c + \sum_{k,l=0}^{p_i} (\xi - \xi_{kl}) X_{kl} \\ &+ \sum_{k,l=0}^{p_i, q_i} (\xi_{kl} - \xi_k^r) X_{kl} + \sum_{k,l=0}^{q_i, p_i} (\xi_{kl} - \xi_l^c) X_{kl}, \quad i = 0, 1, 2, \cdots. \end{aligned} \tag{1.5}$$

Then each $x_i \in rc$, and it is easy to verify that p_i, q_i, $(q_i \geqq p_i \to \infty)$, can be so chosen that $\lim_i x_i = x$.

2. **The general form of linear functionals in** rc. The following fact will be needed in the proof of Lemma 1 in §3.

Every linear functional $f(x)$ *defined in* rc *is of the form*

$$f(x) = C\xi + \sum_k C_k^r \xi_k^r + \sum_l C_l^c \xi_l^c + \sum_{k,l} C_{kl} \xi_{kl}, \tag{2.1}$$

and

---

[3] Compare with Banach, op. cit., p. 90.

[4] See Banach, op. cit., p. 47, Theorem 10. The proof given is easily extended to the system (1.3).

[5] Compare with Banach, op. cit., p. 90.

[6] Compare with Banach, op. cit., p. 90.

$$\|f\| = |C| + \sum_k |C_k^r| + \sum_l |C_l^c| + \sum_{k,l} |C_{kl}|. \tag{2.2}$$

PROOF. Let $f(x)$ be an arbitrary linear functional defined in rc, and set $f(X) \equiv A$, $f(X_k^r) \equiv A_k^r$, $f(X_l^c) \equiv A_l^c$, $f(X_{kl}) \equiv A_{kl}$, $(k, l = 0, 1, 2, \cdots)$. We show first that each of the series $\sum_k |A_k^r|$, $\sum_l |A_l^c|$, $\sum_{k,l} |A_{kl}|$ is convergent.

Let $x \in rcrn$ be arbitrary and set $x_{ij} \equiv \sum_{k,l=0}^{i,j} \xi_{kl} X_{kl}$, $(i, j = 0, 1, 2, \cdots)$. Then each $x_{ij} \in rcrn$ and since $x_{ij} \to x$ as $i, j \to \infty$, it follows that $f(x_{ij}) = \sum_{k,l=0}^{i,j} \xi_{kl} A_{kl}$ tends to a limit, namely $f(x)$, as $i, j \to \infty$. In order that this be true for every $x \in rcrn$ it is necessary[7] that $\sum_{k,l} |A_{kl}| < \infty$.

On the other hand let $\{\xi_k\}$ be an arbitrary null sequence, and let $x \equiv \{\xi_{kl}\}$ be such that for each fixed $k = 0, 1, 2, \cdots$ we have $\xi_{kl} = \xi_k$ for $l = 0, 1, 2, \cdots$. If we set $x_i \equiv \sum_{k=0}^{i} \xi_k X_k^r$, $(i = 0, 1, 2, \cdots)$, then each $x_i \in rc$ and $x_i \to x$ as $i \to \infty$. Since $f(x_i) = \sum_{k=0}^{i} \xi_k A_k^r$, it follows that the series $\sum_k \xi_k A_k^r$ is convergent for every null sequence $\{\xi_k\}$. Consequently, the condition $\sum_k |A_k^r| < \infty$ must be satisfied,[8] and it is evident that the symmetrical condition $\sum_l |A_l^c| < \infty$ must also hold.

Now let $x \in rc$ be chosen arbitrarily, and let x_i be defined as in (1.5) so that $x_i \to x$. Then $f(x_i) \to f(x)$, and from (1.5) we find that

$$f(x) = \xi A + \sum_k (\xi_k^r - \xi) A_k^r + \sum_l (\xi_l^c - \xi) A_l^c + \sum_{k,l} (\xi + \xi_{kl} - \xi_k^r - \xi_l^c) A_{kl}.$$

On rearranging the latter expression and setting

$$C \equiv A - \sum_k A_k^r - \sum_l A_l^c + \sum_{k,l} A_{kl}, \qquad C_k^r \equiv A_k^r - \sum_l A_{kl},$$

$$C_l^c \equiv A_l^c - \sum_k A_{kl}, \qquad C_{kl} \equiv A_{kl},$$

$f(x)$ reduces to the form (2.1) as was to be shown.

Regarding the evaluation of $\|f\|$, it is clear that $\|f\|$ does not exceed the value given in (2.2). The converse statement can be readily shown to follow from the relation $|f(x_n)| \leqq \|f\|$, wherein $x_n \equiv \{\xi_{kl}^n\}$ for each $n = 0, 1, 2, \cdots$ is defined as follows:[9] $\xi_{kl}^n \equiv \operatorname{sgn} C_{kl}$ for $0 \leqq k, l \leqq n$; $\equiv \operatorname{sgn} C_k^r$ for $0 \leqq k \leqq n$, $n < l < \infty$; $\equiv \operatorname{sgn} C_l^c$ for $0 \leqq l \leqq n$, $n < k < \infty$;

[7] See Hamilton, loc. cit., p. 41, #1.

[8] See Banach, op. cit., p. 86. Merely replace (c) by (c_0) in the discussion at the top of the page.

[9] To insure $\|x_n\| = 1$ for every n, let $\operatorname{sgn} \alpha \equiv 1$ if $\alpha \geqq 0$; $\equiv -1$ if $\alpha < 0$.

$\equiv \operatorname{sgn} C$ for $n<k$, $l<\infty$. Inasmuch as we make no use of (2.2), the remaining details may be omitted.

3. **The analogues of Banach's results.** We come next to the following four lemmas. Since these bear a direct analogy to those of Banach, both in statement and method of proof, we shall give here only the proof of the first lemma. The second one is a straightforward corollary of the first, and the reader who is familiar with the original lemmas will encounter no difficulty in supplying the remaining two proofs.[10]

LEMMA 1. *Let A be a completely regular method and let $y_0 \equiv \{\eta_{ij}^0\}$ be an arbitrary element of rc. If the conditions* (1.4) *always imply* $\sum_{i,j}\alpha_{ij}\eta_{ij}^0 = 0$, *then corresponding to each $\epsilon>0$ there exists an $x \in rc$ such that*

$$|A_{ij}(x) - \eta_{ij}^0| < \epsilon, \qquad i, j = 0, 1, 2, \cdots. \tag{3.1}$$

PROOF. Let G denote the subset of rc of all sequences of the form $\{\eta_{ij}\} \equiv \{A_{ij}(x)\}$ for x ranging over rc. Evidently G is a linear set.

Let us assume that for some particular $\epsilon>0$, the condition (3.1) cannot be satisfied. This is equivalent to assuming that y_0 is not a limit element of G. In such a case there must exist a linear functional $F(y)$ defined in rc such that[11] $F(y_0)=1$ and $F(y)=0$ for every $y \in G$. In view of §2 we are therefore led to relations of the following sort:

$$C\eta^0 + \sum_i C_i^r \eta_i^{0r} + \sum_j C_j^c \eta_j^{0c} + \sum_{i,j} C_{ij}\eta_{ij}^0 = 1, \tag{3.2}$$

$$C\eta + \sum_i C_i^r \eta_i^r + \sum_j C_j^c \eta_j^c + \sum_{i,j} C_{ij}\eta_{ij} = 0 \quad \text{for all} \quad \{\eta_{ij}\} \in G, \tag{3.3}$$

wherein $\sum_{i,j}|C_{ij}| < \infty$. From the definition of G and the fact that A is completely regular, (3.3) may be written

$$C\xi + \sum_i C_i^r \xi_i^r + \sum_j C_j^c \xi_j^c + \sum_{i,j} C_{ij}A_{ij}(x) = 0 \quad \text{for all} \quad x \in rc. \tag{3.4}$$

Now by virtue of (1.2) we have $\sum_{i,j,k,l=0}^{n}|C_{ij}a_{ijkl}\xi_{kl}| \leqq M\cdot\|x\| \cdot\sum_{i,j}|C_{ij}| < \infty$ for each $n=0, 1, 2, \cdots$ and each $x \equiv \{\xi_{kl}\} \in rc$. Consequently, (3.4) may be expressed in the form

$$\Phi(x) \equiv C\xi + \sum_i C_i^r \xi_i^r + \sum_j C_j^c \xi_j^c + \sum_{k,l} \xi_{kl} \sum_{i,j} C_{ij}a_{ijkl} = 0,$$

valid for all $x \in rc$. Allowing x to take in turn the values X_{kl}, X_k^r, X_l^c,

[10] In this connection, however, see the remark in §1.

[11] Banach, op. cit., p. 57, §3, Lemma.

X, $(k, l=0, 1, 2, \cdots)$, in $\Phi(x)$, we find successively the following relations:

$$\sum_{i,j} C_{ij}a_{ijkl} = 0, \qquad k, l = 0, 1, 2, \cdots, \tag{3.5}$$

and $C_i'=C_j'=C=0$, $(i, j=0, 1, 2, \cdots)$. Equation (3.2) then reduces to $\sum_{i,j}C_{ij}\eta_{ij}^0=1$, and this in view of (3.5) contradicts the hypothesis. The proof is therefore complete.

LEMMA 2. *If the method A is completely regular and of type M, then corresponding to each $\{\eta_{ij}^0\}$ ε rc and each $\epsilon>0$, there exists an x ε rc satisfying condition* (3.1).

LEMMA 3. *If x_0 is a bounded sequence transformed into rc by a completely regular method A, then corresponding to each $\epsilon>0$ there exists an x ε rc such that*

$$|A_{ij}(x) - A_{ij}(x_0)| < \epsilon, \qquad i, j = 0, 1, 2, \cdots. \tag{3.6}$$

LEMMA 4. *Let x_0 be any sequence transformed into rc by a completely regular and reversible method A. If corresponding to each $\epsilon>0$ there exists an x ε rc satisfying* (3.6), *then the sequence x_0 is summable to the same principal limit by every method B which transforms rc into c regularly and which transforms into c every sequence that A transforms into rc. Moreover, if B is completely regular and transforms into rc every sequence that A does, then the row and column limits assigned by B to x_0 will be respectively equal to those assigned by A.*

Finally, the preceding lemmas may be combined to yield the following counterpart of Banach's theorems.

THEOREM. *Let the method A be perfect* [*completely regular and reversible*], *and let the method B transform rc into c regularly and transform into c every sequence that A transforms into rc. Then every sequence* [*bounded sequence*] *x_0 transformed into rc by A is transformed into c by B with the same principal limit. Moreover, if B is completely regular and transforms into rc every sequence that A does, then the row and column limits assigned by B to x_0 will be respectively equal to those assigned by A.*

MICHIGAN STATE COLLEGE

NOTE ON THE COMPUTATION OF THE DIFFERENCES OF THE Si(x), Ci(x), Ei(x) AND $-\mathrm{Ei}(-x)$ FUNCTIONS

MILTON ABRAMOWITZ

In a recent article,[1] Lowan reached the conclusion that the second differences of the functions Si(x), Ci(x) and Ei(x) for values of x from 0.1 to 2 at intervals of 0.0001 may be approximated to 12 decimal places by certain expressions in closed form, which may be computed with the aid of tabulated functions.

The above functions and also $-\mathrm{Ei}(-x)$ were computed to 12 decimal places by adding an appropriate number of terms in their power series expansion.[2] In order to check the accuracy of the computations, the second differences of the computed values were required to agree to 12 decimal places with the values of the approximate expressions mentioned above.

The functions Ci(x), Ei(x), $-\mathrm{Ei}(-x)$ have a logarithmic singularity at the origin. For this reason Lowan's conclusion could not be extended to the interval $0<x<0.1$.

If $\phi(x)$ represents any of the functions $\mathrm{Ci}(x)-\log_e x$, $\mathrm{Ei}(x)-\log_e x$, $-\mathrm{Ei}(-x)+\log_e x$, it will be shown that the second difference of $\phi(x)$ may be approximated by $h^2\phi''(x)$ to 12 decimal places. Let

$$R(x) = [\phi(x+h) - 2\phi(x) + \phi(x-h)] - h^2\phi''(x).$$

Substituting for $\phi(x+h)$ and $\phi(x-h)$ their Taylor expansions, we get

$$R(x) = 2\left[\frac{h^4}{4!}\phi^4(x) + \frac{h^6}{6!}\phi^6(x) + \cdots\right]$$

whence

$$\text{(A)} \qquad |R(x)| < 2\left[\frac{h^4}{4!}\{\phi^4(x)\} + \frac{h^6}{6!}\{\phi^6(x)\} + \cdots\right]$$

where $\{\phi^{2k}(x)\}$ is an upper bound of the modulus of the $2k$th derivative.

---

[1] A. N. Lowan, *On the computation of second differences of the* Si(x), Ei(x), *and* Ci(x) *functions*, this Bulletin, vol. 45 (1939), pp. 583–588. This will be referred to as A.N.L.

[2] This work is done by the New York City Work Projects Administration project for the Computation of Mathematical Tables, O. P. #765–97–3–10, under the sponsorship of Dr. Lyman J. Briggs, Director of the National Bureau of Standards. The author wishes to express his appreciation to the W.P.A. and to the sponsor of this project for permission to publish these results.

Case of the function $\mathrm{Ei}(x)-\log_e x$. From

$$\frac{d}{dx}(\mathrm{Ei}\,(x)-\log_e x)=\frac{e^x-1}{x}=\int_0^\infty e^{x(1-t)}dt-\int_0^\infty e^{xt}dt$$

(see A.N.L.) we get

$$\frac{d^{2k}}{dx^{2k}}(\mathrm{Ei}\,(x)-\log_e x)=\int_0^\infty (1-t)^{2k-1}e^{x(1-t)}dt-\int_0^\infty t^{2k-1}e^{xt}dt,$$

and, in view of (9) and (10) of A.N.L.,

$$\left|\frac{d^{2k}}{dx^{2k}}(\mathrm{Ei}\,(x)-\log_e x)\right|<\frac{e^x}{2k}.$$

Making use of this result (A) yields

$$|R(x)|<2e^x\left\{\frac{h^4}{4!}\,\frac{1}{4}+\frac{h^6}{6!}\,\frac{1}{6}+\cdots\right\}\sim\frac{e^x}{2}\,\frac{h^4}{4!}\,\frac{1}{1-h}.$$

For $h=10^{-4}$ and $0<x<0.1$, the last expression is smaller than 10^{-17}. We thus reach the conclusion that for these values of h and x, the second differences of $\mathrm{Ei}(x)-\log_e x$ may be approximated by

$$h^2\frac{d^2}{dx^2}(\mathrm{Ei}\,(x)-\log_e x).$$

From the series expansion for the function $\mathrm{Ei}(x)-\log_e x$ it is easily seen that the second difference of this function may be approximated to 12 decimal places by

$$h^2\left(\frac{1}{2}+\frac{x}{3}+\frac{1}{4}\,\frac{x^2}{2!}+\frac{1}{5}\,\frac{x^3}{3!}\right).$$

In an entirely similar manner it may be shown that $1/(2k)$ is an upper bound of the derivatives of the functions $-\mathrm{Ei}(-x)+\log_e x$ and $\mathrm{Ci}(x)-\log_e x$, and if $h=10^{-4}$ and $0<x<0.1$, the second difference of the latter functions may be approximated to 12 decimal places by the respective expressions

$$h^2\left(-\frac{1}{2}+\frac{x}{3}-\frac{1}{4}\,\frac{x^2}{2!}+\frac{1}{5}\,\frac{x^3}{3!}\right),\qquad h^2\left(-\frac{1}{2}+\frac{1}{4}\,\frac{x^2}{2!}\right).$$

NEW YORK CITY

DIOPHANTINE EQUATIONS OF DEGREE[1] n

A. A. AUCOIN

In a recent issue of the National Mathematics Magazine,[2] W. V. Parker and the author obtained solutions of the Diophantine equation $F(x_1, \cdots, x_p) = G(y_1, \cdots, y_q)$, where F and G are homogeneous polynomials, with integral coefficients, of degree 3, and F is such that for a set of integers $x_i = a_i$ (not all zero), $\partial F/\partial x_i = 0$, $(i = 1, \cdots, p)$. In this paper the above is extended to functions of degree n. One type, which satisfies the conditions of the main theorem, is also solved by an entirely different method. The solutions obtained are in terms of arbitrary parameters, and they are integral for an integral choice of the parameters.

If $x_i = \alpha_i$, $y_k = \beta_k$ is a solution of the equation $f(x_1, \cdots, x_p) = g(y_1, \cdots, y_q)$, where f and g are homogeneous polynomials, with integral coefficients, of degrees n and m respectively, and there are no integers $s > 1$, α_i', β_k' such that $\alpha_i = s^\lambda \alpha_i'$, $\beta_k = s^\mu \beta_k'$ where λ, μ are relatively prime positive integers such that $\lambda n = \mu m$, then $x_i = \alpha_i$, $y_k = \beta_k$ is said to be a primitive solution. If $x_i = \alpha_i$, $y_k = \beta_k$ is a primitive solution of the above equation, then $x_i = \alpha_i t^\lambda$, $y_k = \beta_k t^\mu$ (derived from this primitive solution), where λ, μ are any positive integers such that $\lambda n = \mu m$, is also a solution. Two solutions are said to be equivalent if they are derived from the same primitive solution.

THEOREM 1. *Let $f(x_1, \cdots, x_p)$, $g(y_1, \cdots, y_q)$ be homogeneous polynomials with integral coefficients, of degrees n and m respectively. Let $a_1, \cdots, a_p$ be integers not all zero such that the partial derivatives of f of all orders less than $n-1$ vanish[3] when $x_i = a_i$. Then every solution in integers x_i, y_k of*

$$f(x_1, \cdots, x_p) = g(y_1, \cdots, y_q), \tag{1}$$

for which

$$\sum_{j=1}^{p} a_j \frac{\partial f}{\partial x_j} \neq 0, \tag{2}$$

is equivalent to one of the infinitude of solutions given by

$$x_i = a_i s t^{\lambda-1} + \alpha_i t^\lambda, \quad y_k = \beta_k t^\mu, \quad i = 1, 2, \cdots, p; \; k = 1, 2, \cdots, q, \tag{3}$$

[1] Presented to the Society, December 2, 1939.

[2] *On cubic Diophantine equations*, vol. 13 (1938), pp. 115–117.

[3] It follows from Euler's theorem that the function itself vanishes for this choice of x_i.

where λ, μ are positive integers such that $\lambda n=\mu m$, α_i and β_k are arbitrary integers,

$$s = g(\beta) - f(\alpha), \qquad t = \sum_{j=1}^{p} a_j \frac{\partial f}{\partial \alpha_j}, \tag{4}$$

and $f(\alpha)=f(\alpha_1, \cdots, \alpha_p)$, $g(\beta)=g(\beta_1, \cdots, \beta_q)$.

PROOF. By Taylor's theorem, if we set $x_i=a_i s+\alpha_i t$,

$$f(x_1, \cdots, x_p) = st^{n-1}\sum_{j=1}^{p} a_j \frac{\partial f}{\partial \alpha_j} + t^n f(\alpha_1, \cdots, \alpha_p).$$

Hence if x_i, y_k have the values given by (3), s and t those given by (4), (1) becomes $st^{n\lambda-1}\sum_{j=1}^{p} a_j \partial f/\partial \alpha_j + t^{n\lambda} f(\alpha) = t^{m\mu} g(\beta)$, and is satisfied identically in the α_i and β_k. Hence (3) is a solution of (1) with s and t given by (4).

Suppose $x_i=\rho_i$, $y_k=\sigma_k$ is any solution of (1) and (2). If we choose $\alpha_i=\rho_i$, $\beta_k=\sigma_k$, we have that $s=0$, and (3) becomes $x_i=\rho_i t^{\lambda}$, $y_k=\sigma_k t^{\mu}$, equivalent to the given solution, since by (2), $t\neq 0$.

If $g\equiv 0$, the theorem still holds, with λ arbitrary.

COROLLARY. *The equation $f(x)=\sum_{j=1}^{p} x_j g_j(y)+g(y)$, where $g_j(y)=g_j(y_1, \cdots, y_q)$ and $g(y)=g(y_1, \cdots, y_q)$ are homogeneous polynomials with integral coefficients of degrees $n-1$ and n, respectively, has solutions, and every solution which is not also a solution of $\sum_{j=1}^{p} a_j[\partial f/\partial x_j - g_j(y)]=0$ is equivalent to one of the infinitude of solutions given by $x_i=a_i s+\alpha_i t$, $y_k=\beta_k t$ where*

$$s = g(\beta) - f(\alpha) - \sum_{j=1}^{p} \alpha_j g_j(\beta), \qquad t = \sum_{j=1}^{p} a_j\left[\frac{\partial f}{\partial \alpha_j} - g_j(\beta)\right].$$

One function of interest which satisfies the hypothesis of Theorem 1 is the function $D(x)=|a_{ij}x_{ij}|$, a determinant of order n with a_{ij} integral such that not all the a's in any row or column are zero. For this function not all the x_{ij} need be distinct. If there is any x_{ij}, say x_{pq}, which occurs only once in D, we may make the choice $x_{pq}=1$, $x_{ij}=0$ otherwise; then all the partial derivatives of all orders less than $n-1$ vanish.

In the solution the form of the expression is the same for every element except x_{pq}. This fact is illustrated by the equation $D(x)=g(y)$, the solution of which is $x_{pq}=st^{\lambda-1}+\alpha_{pq}t^{\lambda}$, $x_{ij}=\alpha_{ij}t^{\lambda}$, $(i\neq p,\ j\neq q)$, $y_k=\beta_k t^{\mu}$ where $s=g(\beta)-D(\alpha)$, $t=D'(\alpha)$, $D'(\alpha)$ being $D(\alpha)$ with a_{pq} the element in the pth row and qth column, and the other elements in the qth column zero.

It is not necessary, in some cases, that there be a unique element x_{pq}. If $a_{ij}=1$, for example, D may be the circulant. In this case we make the choice $x_{ij}=1$.

Another function of interest which also satisfies the condition of Theorem 1 is $P(x)=\prod_{i=1}^{n}\sum_{j=1}^{n}a_{ij}x_j$, where a_{ij} are integral, and the determinant $A=|a_{ij}|\neq 0$. For we may choose x_j, integral, so that $n-1$ of the above factors vanish and hence for this choice of x_j all partial derivatives of all orders less than $n-1$ vanish.

The next equation satisfies the hypothesis of Theorem 1, but will be solved by an entirely different method. This is given in the following theorem:

THEOREM 2. *The equation*

$$P(x) = g(y), \tag{5}$$

where $P(x)$ is given above and $g(y)$ is given in Theorem 1, *has solutions, and every solution which is not also a solution of*

$$\prod_{i=1}^{n-1}\sum_{j=1}^{n}a_{ij}x_j = 0 \tag{6}$$

is equivalent to one of the infinitude of solutions given by

$$\begin{aligned} x_j &= t^{\lambda}(A_{nn})^{\lambda-1}A_{nn}^{(j)} + st^{\lambda-1}(A_{nn})^{\lambda-1}A_{nj}, \quad j=1,\cdots,n-1,\\ x_n &= st^{\lambda-1}(A_{nn})^{\lambda}, \qquad y_k = t^{\mu}(A_{nn})^{\mu}\beta_k, \end{aligned} \tag{7}$$

where A_{ij} is the cofactor of a_{ij} in[4] *A; $A_{nn}^{(j)}$ is the determinant obtained from A_{nn} by replacing the jth column by $\alpha_1, \alpha_2, \cdots, \alpha_{n-1}$; s and t are given by*

$$s = A_{nn}g(\beta) - \prod_{i=1}^{n-1}\alpha_i\sum_{j=1}^{n}a_{nj}A_{nn}^{(j)}, \qquad t = A\prod_{i=1}^{n-1}\alpha_i, \tag{8}$$

λ, μ are relatively prime positive integers such that $\lambda n=\mu m$, and the α's and β's are arbitrary integers.

PROOF. Set

$$\sum_{j=1}^{n}a_{ij}x_j = t^{\lambda}(A_{nn})^{\lambda}\alpha_i, \qquad i=1,\cdots,n-1. \tag{9}$$

If we let $x_n=st^{\lambda-1}(A_{nn})^{\lambda}$, we may write equation (9) in the form

[4] Since $A\neq 0$, there is a minor of order $n-1$ which does not vanish. Without loss of generality, we may choose the notation so that $A_{nn}\neq 0$.

$\sum_{j=1}^{n-1} a_{ij}x_j = t^{\lambda}(A_{nn})^{\lambda}\alpha_i - st^{\lambda-1}a_{in}(A_{nn})^{\lambda}$. Solving this system of equations we get

$$x_j = t^{\lambda}(A_{nn})^{\lambda-1}A_{nn}^{(j)} + st^{\lambda-1}(A_{nn})^{\lambda-1}A_{nj}, \qquad j = 1, \cdots, n-1. \tag{10}$$

It follows then that

$$\sum_{j=1}^{n} a_{nj}x_j = t^{\lambda-1}(A_{nn})^{\lambda-1}\left[t\sum_{j=1}^{n-1} a_{nj}A_{nn}^{(j)} + sA\right]. \tag{11}$$

If we let $y_k = t^{\mu}(A_{nn})^{\mu}\beta_k$, then by (9) and (11), (5) becomes

$$t^{n\lambda-1}(A_{nn})^{n\lambda-1}\prod_{i=1}^{n-1}\alpha_i\left[t\sum_{j=1}^{n-1} a_{nj}A_{nn}^{(j)} + sA\right] = t^{m\mu}(A_{nn})^{m\mu}g(\beta), \tag{12}$$

and since $\lambda n = \mu m$, (12) is identically satisfied in the α's and β's if s and t are given by (8). Hence (7) forms a solution of (5) with s and t given by (8).

Suppose now that $x_j = \rho_j$, $y_k = \sigma_k$ is any solution of (5). If we choose $\alpha_i = \sum_{j=1}^{n} a_{ij}\rho_j$, $\beta_k = \sigma_k$, we have[5]

$$\begin{aligned}
t &= A\prod_{i=1}^{n-1}\sum_{j=1}^{n} a_{ij}\rho_j, \\
s &= A_{nn}g(\sigma) - \prod_{i=1}^{n-1}\sum_{j=1}^{n} a_{ij}\rho_j\sum_{k=1}^{n-1} a_{nk}[\rho_k A_{nn} - \rho_n A_{nk}] \\
&= A_{nn}\prod_{i=1}^{n}\sum_{j=1}^{n} a_{ij}\rho_j - \prod_{i=1}^{n-1}\sum_{j=1}^{n} a_{ij}\rho_j\sum_{k=1}^{n-1} a_{nk}[\rho_k A_{nn} - \rho_n A_{nk}] \\
&= \prod_{i=1}^{n-1}\sum_{j=1}^{n} a_{ij}\rho_j\left[A_{nn}\sum_{j=1}^{n} a_{nj}\rho_j - A_{nn}\sum_{j=1}^{n-1} a_{nj}\rho_j + \rho_n\sum_{j=1}^{n-1} a_{nj}A_{nj}\right] \\
&= \rho_n A\prod_{i=1}^{n-1}\sum_{j=1}^{n} a_{ij}\rho_j = \rho_n t.
\end{aligned}$$

Hence

$$\begin{aligned}
x_j &= t^{\lambda}(A_{nn})^{\lambda-1}[\rho_j A_{nn} - \rho_n A_{nj}] + (A_{nn})^{\lambda-1}t^{\lambda}\rho_n A_{nj} = (tA_{nn})^{\lambda}\rho_j, \\
x_n &= (tA_{nn})^{\lambda}\rho_n, \qquad\qquad y_k = (tA_{nn})^{\mu}\sigma_k,
\end{aligned}$$

which is equivalent to the given solution provided $x_j = \rho_j$, $y_k = \sigma_k$ is not a solution of (6). We may find, however, values of x_j which satisfy (6) and these values, together with $y_k = 0$, afford additional solutions of (5).

[5] $A_{nn}^{(j)}$ becomes $\rho_j A_{nn} - \rho_n A_{nj}$ when α_i is replaced by $\sum_{j=1}^{n} a_{ij}\rho_j$.

By the above method we may show the following consequence.

COROLLARY. *The equation* $P(x)=\sum_{j=1}^{n}x_jg_j(y)+g(y)$, *where* g_j *and* g *are the functions of the corollary to Theorem* 1, *has solutions, and every solution not also a solution of* $\sum_{j=1}^{n}A_{nj}g_j(y)-A\prod_{i=1}^{n-1}\sum_{j=1}^{n}a_{ij}x_j=0$ *is given by* $x_j=A_{nn}^{(j)}t+A_{nj}s$, $(j=1,\cdots,n-1)$, $x_n=A_{nn}s$, $y_k=A_{nn}\beta_k t$, *where*

$$s=\prod_{i=1}^{n-1}\alpha_i\sum_{j=1}^{n}a_{nj}A_{nn}^{(j)}-\sum_{j=1}^{n-1}A_{nn}^{(j)}g_j(\beta)-A_{nn}g(\beta),$$

$$t=\sum_{j=1}^{n}A_{nj}g_j(\beta)-A\prod_{i=1}^{n-1}\alpha_i.$$

The final theorem treats an equation which satisfies the hypothesis of Theorem 1, but is reduced to an equivalent problem and then solved.

THEOREM 3. *The equation*

$$f(x)\sum_{j=1}^{p}d_jx_j=R(y), \tag{13}$$

where $f(x)$ *satisfies the conditions of Theorem* 1, *and* $R(y)=R(y_1,\cdots,y_q)$ *is a homogeneous polynomial with integral coefficients of degree* $n-1$, *has solutions; and every solution which is not also a solution of*

$$f(x)\sum_{j=1}^{p}a_j\frac{\partial f}{\partial x_j}\left[\sum_{j=1}^{p}a_j\frac{\partial f}{\partial x_j}\sum_{k=1}^{p}d_kx_k-f(x)\sum_{j=1}^{p}d_ja_j\right]=0 \tag{14}$$

is equivalent to one of the infinitude of solutions given by

$$x_i=a_is+\alpha_it,\qquad y_k=\beta_kt \tag{15}$$

where

$$\begin{aligned}s&=A^{n-1}[\lambda(AD-BC)]^{n-2}[D\lambda^2-BR(\mu)],\\ t&=A^{n-1}[\lambda(AD-BC)]^{n-2}[AR(\mu)-C\lambda^2],\\ \beta_k&=A^2[\lambda(AD-BC)]^2\mu_k,\end{aligned} \tag{16}$$

and $A=\sum_{j=1}^{p}a_j\partial f/\partial\alpha_j$, $B=f(\alpha)$, $C=\sum_{j=1}^{p}d_ja_j$, $D=\sum_{j=1}^{p}d_j\alpha_j$, *the* α_j, λ, μ_k *being arbitrary integers.*

PROOF. If we let x_i, y_k have the values given by (15), (13) becomes, after dividing out the factor[6] t^{n-1},

$$(As+Bt)(Cs+Dt)=R(\beta). \tag{17}$$

[6] It will be shown later that $t\neq 0$.

By Theorem 2, the solutions of the equation (17) are given by (16).

If $x_i=\rho_i$, $y_k=\sigma_k$ is any solution of (13) and we choose $\alpha_i=\rho_i$, $\mu_k=\sigma_k$, $\lambda=f(\rho)$, we have that $s=0$ and the solution becomes $x_i=\rho_i K^{n-1}$, $y_k=\sigma_k K^{n+1}$, where $K=A\lambda(AD-BC)$, which is equivalent to the given solution provided $K\neq 0$; that is, provided $x_i=\rho_i$, $y_k=\sigma_k$ is not a solution of (14). It will be noted that if $K\neq 0$, then $t\neq 0$.

LOUISIANA STATE UNIVERSITY

A MULTIPLE NULL-CORRESPONDENCE AND A SPACE CREMONA INVOLUTION OF ORDER $2n-1$[1]

EDWIN J. PURCELL

PART I. A NULL-SYSTEM $(1, mn, m+n)$ BETWEEN THE PLANES AND POINTS OF SPACE $(m, n=1, 2, 3, \cdots)$

1. **Introduction.** Consider a curve δ_m of order m having $m-1$ points in common with a straight line d, and a curve δ_n' of order n having $n-1$ points in common with a straight line d', $(m, n=1, 2, 3, \cdots)$. It is assumed for the present that neither δ_m nor d intersects either δ_n' or d'.

In general, through any point P of space there passes one ray ρ which intersects δ_m once and d once, and one ray ρ' which intersects δ_n' once and d' once; ρ and ρ' determine a plane π, the null-plane of P. Conversely, a plane π determines m rays ρ_i and n rays ρ_j' lying in it which intersect, a ray ρ with a ray ρ', in mn points, the null-points of the plane π.

Any point α in general position determines a ray ρ. As α describes a line l, the plane π of ρ and l contains n rays ρ', which intersect l in n points β; conversely, any point β on l determines a ray ρ' which determines with l the plane π, and π contains m rays ρ which intersect l in m points α—one being the original α. Thus an (m, n) correspondence is set up among the points of l with valence zero; there are $m+n$ coincidences and therefore $m+n$ points on any line l whose null-planes contain l.

2. **Planes whose null-points behave peculiarly.** We can obtain the last result by another method; this will yield additional information about planes whose null-points behave peculiarly.

Let a plane π turn about a line l as axis. A ruled surface will be generated by the m rays ρ_i lying in π. This surface is of order $m+1$; δ_m is a onefold curve on the surface and d is an m-fold line. Another

[1] Presented to the Society, December 2, 1939.

ruled surface will be generated in this manner by the rays ρ_j' lying in π; its order is $n+1$, δ_n' is a onefold curve and d' is an n-fold line on this surface. The curve of intersection of these two surfaces is of order $(m+1)(n+1)$ and consists of l and a twisted curve k_{mn+m+n} of order $(m+1)(n+1)-1=mn+m+n$. *This k_{mn+m+n} is the locus of the null-points of all planes π through l.*

Since a plane π meets this in mn points outside l, k_{mn+m+n} must intersect l in $m+n$ points through each of which a ray ρ and a ray ρ' pass which are coplanar with l. Call such a point on l, P. The plane $\rho\rho'$ is the null-plane of P and has $mn-1$ null-points outside l, and it follows that plane $\rho\rho'$ is tangent to k_{mn+m+n} at P. *The null-planes of the $m+n$ points of intersection of k_{mn+m+n} with l are tangent planes of k_{mn+m+n} at these points.*

The line d, an m-fold line on the first of the two surfaces described above, intersects the second surface in $n+1$ points, which are m-fold points on the first surface. The line d' intersects the first of the two surfaces in $m+1$ points which are n-fold points on the second surface. These points all lie on k_{mn+m+n} and the $m+1$ are n-fold points of it and $n+1$ are m-fold points of it. *k_{mn+m+n} has $m+1$ n-fold points on d' and $n+1$ m-fold points on d.*

δ_m has no actual double points or other multiple points. It is, however, rational and has $(m-1)(m-2)/2$ apparent double points and its rank is $r=m(m-1)-(m-1)(m-2)=2(m-1)$; that is, the order of its developable surface is $2(m-1)$. Similarly, the order of the developable surface of δ_n' is $2(n-1)$. The line l will intersect $2(m-1)$ tangents of δ_m and $2(n-1)$ tangents of δ_n'. In the plane π through l and a tangent line t of the first group, two rays ρ coincide in the line which joins the point of tangency of t with the intersection of d and π. Of the mn null-points in the plane π, n lie on each of the other $m-2$ rays ρ, and $2n$ fall two and two together on the coinciding rays; *in these points k_{mn+m+n} is tangent to the plane of l and t and the number of these planes is $2(m+n-2)$.*

From the discussion of this section we have the following conclusions:

(1) *The planes, m of whose null points coincide with a point of d, envelope a surface of class $n+1$; and the planes, n of whose null points coincide with a point of d, envelope a surface of class $m+1$.*

(2) *The planes, $2n$ of whose null-points coincide two and two on a ray ρ, envelope a surface of class $2(m-1)$, n of the remaining null-points lying on each of the other $m-2$ rays ρ; the planes, $2m$ of whose null-points coincide two and two on a ray ρ', envelope a surface of class*

$2(n-1)$, m of the remaining null-points lying on each of the other $n-2$ rays ρ'.

Consider a plane π through l, whose intersection with d is also an intersection with δ_m. Call this common point of d and δ_m, Δ. Then the rays ρ_i lying in π will be the $m-1$ lines joining Δ to the $m-1$ points of intersection of δ_m and π, not lying on d, and the line λ joining Δ to the intersection of l and the plane of d and the tangent line to δ_m at Δ. This line λ will be the limiting position of a ray ρ as a plane revolves about l into the position of π.

In the osculating planes of δ_m and δ_n', three rays coincide. Therefore, *in the osculating planes of δ_m, $3n$ of the null-points coincide three and three on the triple ray; in the osculating planes of δ_n', $3m$ of the null-points coincide three and three on the triple ray.*

3. **Points whose null-planes behave peculiarly.** Consider a point P on d. The point P determines one ρ'. Any plane π through ρ' determines m rays ρ through P. Therefore π counts m times as null-plane of P. Conversely, for every plane through ρ' there fall m null-points together at P. The surface of class $n+1$ mentioned in §2 must have the planes π as tangent planes. *This surface is a ruled surface consisting of rays ρ' which intersect d, and conversely.* Call this surface Σ.

The surface formed by rays ρ' which intersect a general straight line l is (§2) of order $n+1$, and d intersects this surface in $n+1$ points. Thus there are n rays ρ' which intersect d and also an arbitrary line l. Therefore the surface Σ is of degree $n+1$. The line d is a onefold directrix on Σ_{n+1} and d' is an n-fold directrix; for, the n-ic cone of δ_n' projected from a point of d' will intersect d in n points. *The locus of points whose null-planes have m null-points coinciding is Σ_{n+1}.*

Similarly, the ruled surface Σ_{m+1}' of order $m+1$, consisting of rays ρ that intersect d', is the *locus of points whose null-planes have n null-points coinciding.*

Now Σ_{n+1} and Σ_{m+1}' have $mn+1$ generators in common. For the congruence of rays ρ has the characteristic $(1, m)$ and the congruence of rays ρ' has the characteristic $(1, n)$ so that, from Halphen's theorem,[2] there are $1\cdot 1+m\cdot n=mn+1$ common rays.

Since both rays ρ and ρ' through any point on one of these $mn+1$ common rays coincide, any plane through the ray can be taken as null-plane of the point. *Every plane of the pencil through any one of the $mn+1$ common rays has m null-points coinciding on d and n null-points coinciding on d'.*

[2] C. M. Jessop, *A Treatise on the Line Complex*, 1903, p. 259.

The intersection of Σ_{n+1} and Σ'_{m+1} is of degree $(n+1)(m+1)$. Since d' was shown to be an n-fold line on Σ_{n+1} and is clearly a onefold line on Σ'_{m+1}, d' therefore counts n times in the intersection of these two surfaces. Similarly d counts m times in the intersection. Each of the $mn+1$ common rays of the two congruences counts once in the intersection. The parts just enumerated have total degree $n+m+mn+1 = (n+1)(m+1)$. Therefore, *the locus of points whose null-planes have m null-points coinciding in one point and n null-points coinciding in another consists of the lines d and d' and the mn+1 common rays of the two congruences.*

Now consider a plane containing d; let it intersect d' in D' and δ'_n in n points N_i. *Every point of the n lines $D'N_i$ is a null-point of this plane*—similarly for planes through d'.

Let point P be on δ_m but not on d. One ρ' is determined but every line from P to d will be a ρ. Therefore, *any point of δ_m or δ'_n not also a point of d or d' has the pencil of planes through the ray of the opposite congruence as null-planes.*

Part II. A space cremona involution of order $2n-1$ (n any integer)

4. **Definition.** Not every skew curve of order n has a secant meeting it in $n-1$ points, and some have only one such secant, but there are also skew curves of order n that have two $(n-1)$-secant lines. In such case they lie on a quadric surface and have a singly infinite system of such secants. The two selected must be two generators of the same regulus.

Consider a fixed twisted curve δ_n of order n having $n-1$ points in common with a fixed line d and $n-1$ points in common with another fixed line d'. This construction occurs when the two twisted curves δ'_n and δ_m in Part I are identical but lines d and d' remain skew to each other.

A general point P determines a unique line intersecting δ_n once, at A, and d once, at D, and a unique line intersecting δ_n once, at B, and d' once, at D'. We define P', the correspondent of P, to be the intersection of lines AD' and BD. It is an involution.

5. **Equations.** Let d be $x_1=0$, $x_2=0$, and d' be $x_3=0$, $x_4=0$, and the parametric equations of δ_n be

$$x_1 = (as+bt)\prod_1^{n-1}(t_is-s_it), \qquad x_2 = (cs+dt)\prod_1^{n-1}(t_is-s_it),$$

$$x_3 = (es+ft)\prod_n^{2n-2}(t_is-s_it), \qquad x_4 = (gs+ht)\prod_n^{2n-2}(t_is-s_it),$$

where (s_i, t_i), $(i=1, 2, \cdots, n-1)$, are values of the parameter at the $n-1$ points of δ_n on d, and for $i=n, n+1, \cdots, 2n-2$ are values of the parameter at the $n-1$ points of δ_n on d'. Then the equations of the involution are

$$x_1' = (ad-bc)\{(ah-bg)x_3-(af-be)x_4\}\prod_1^{n-1}\alpha_i\prod_1^{n-1}\beta_i,$$

$$x_2' = (ad-bc)\{(ch-dg)x_3-(cf-de)x_4\}\prod_1^{n-1}\alpha_i\prod_1^{n-1}\beta_i,$$

$$x_3' = (fg-eh)\{(cf-de)x_1-(af-be)x_2\}\prod_n^{2n-2}\alpha_i\prod_n^{2n-2}\beta_i,$$

$$x_4' = (fg-eh)\{(ch-dg)x_1-(ah-bg)x_2\}\prod_n^{2n-2}\alpha_i\prod_n^{2n-2}\beta_i,$$

where $\alpha_i \equiv (t_id+s_ic)x_1-(t_ib+s_ia)x_2$ and $\beta_i \equiv (t_ih+s_ig)x_3-(t_ib+s_ie)x_4$. It is of order $2n-1$, n any integer.

6. **The fundamental system.** Line d is an $(n-1)$-fold fundamental line of simple contact. The $n-1$ fixed tangent planes through d are $\alpha_i=0$, $(i=1, 2, \cdots, n-1)$. The line d is an F-line of the first species whose principal surface consists in the $n-1$ planes $\beta_i=0$, $(i=1, 2, \cdots, n-1)$.

Line d' is an $(n-1)$-fold F-line of simple contact. The $n-1$ fixed tangent planes through d' are $\beta_i=0$, $(i=n, n+1, \cdots, 2n-2)$. d' is an F-line of the first species whose P-surface is $\prod_n^{2n-2}\alpha_i=0$.

Points Δ_i, $(i=1, 2, \cdots, n-1)$, intersections of d with δ_n whose parameters on δ_n are (s_i, t_i), and points Δ_i', $(i=n, n+1, \cdots, 2n-2)$, intersections of d' with δ_n, are isolated n-fold F-points whose P-surfaces are, respectively, the above mentioned fixed tangent planes $\alpha_i=0$, $(i=1, 2, \cdots, n-1)$, and $\beta_i=0$, $(i=n, n+1, \cdots, 2n-2)$.

The $(n-1)^2$ lines, each joining a Δ_i to a Δ_i', are simple F-lines without contact. They are F-lines of the second species.

The $(n-1)^2$ lines of intersection of the fixed tangent planes through d with the fixed tangent planes through d' are simple F-lines without contact. They are F-lines of the second species.

7. **Invariant locus.** Every point of the curve δ_n is invariant. Every line that intersects d, d', and δ_n, each once, goes over into itself although it is not pointwise invariant. The locus of these lines is the quadric surface on which d, d', and δ_n lie.

8. **Intersection of two homaloids.** Since they are surfaces of order

$2n-1$, two homaloids intersect in a space curve of order $(2n-1)^2$.

The fixed part of this curve consists in the lines d and d', each counting $n(n-1)$ times, the $(n-1)^2$ lines joining the isolated n-fold F-points of d with those of d', each counting once, and the $(n-1)^2$ lines of intersection of the fixed tangent planes through d with those through d', each counting once. The order of this fixed part is $2n(n-1)+2(n-1)^2$.

The variable part of the curve of intersection is of order $2n-1$ and corresponds to the line of intersection of the two general planes which go over into the pair of homaloids.

UNIVERSITY OF ARIZONA

SYSTEMS OF ORTHOGONAL POLYNOMIALS ON CERTAIN ALGEBRAIC CURVES[1]

FULTON KOEHLER

1. **Introduction.** A system of orthogonal polynomials in two real variables x and y can be defined with respect to any domain of integration in the xy plane and any nonnegative weight function which has a positive integral over that domain. Although a general study of the formal properties of such systems has been made,[2] the corresponding problem of convergence has been adequately dealt with only for cases which reduce almost immediately to familiar problems in a single variable. The purpose of this paper is to give more substantial examples of systems for which it is possible, by special methods, to present an account of convergence which is fully comparable, at least in some of its main features, with such highly developed theories as those of Fourier and Legendre series.

The proofs, as in the case of series of orthogonal polynomials in one variable with a more or less general weight function, can be made to depend on properties of boundedness. The method used here for obtaining these properties is to establish relationships between the systems in two variables considered and systems in one variable whose properties are well known.

The domain of integration to which primary consideration is to be given is the perimeter of the square whose sides are segments of the lines $x=\pm 1$, $y=\pm 1$. Two other domains will be dealt with briefly in a concluding section.

The square contour will be denoted by the letter C. We shall take the following sequence as the basis for the construction of a set of orthogonal polynomials on C: $1,\ x+y,\ x-y,\ xy,\ x^2+y^2,\ x^2-y^2,\ x^2y+xy^2,\ x^2y-xy^2,\ x^3+y^3,\ x^3-y^3,\ \cdots,\ x^{n-1}y+xy^{n-1},\ x^{n-1}y-xy^{n-1},\ x^n+y^n,\ x^n-y^n,\ \cdots$. The terms of this sequence have the property that any finite number of them are linearly independent on C; and also, by means of the identity $x^2y^2-x^2-y^2+1\equiv 0$ which holds everywhere on C, any polynomial in x and y can be expressed on C as a linear combination of them. If $\rho(x, y)$ is a function which is positive almost everywhere on C, we can multiply each member of the sequence by $\rho^{1/2}$ and apply the Schmidt process of orthogonalization,

[1] Presented to the Society, April 14, 1939.

[2] D. Jackson, (1) *Formal properties of orthogonal polynomials in two variables*, Duke Mathematical Journal, vol. 2 (1936), pp. 423–434. (2) *Orthogonal polynomials on a plane curve*, Duke Mathematical Journal, vol. 3 (1937), pp. 228–236.

thus obtaining an orthogonal system of polynomials with respect to the weight function $\rho(x, y)$. Let these polynomials be normalized and given positive leading coefficients, where the leading coefficient of a polynomial is defined as being the coefficient of the last member of the above sequence which appears in it. We shall denote the polynomials thus uniquely defined by $p_{nk}(x, y)$, where n is the degree of the polynomial, and k is 1, 2, 3, or 4 according as the leading term is the first, second, third, or fourth term of the nth degree in the above sequence.

2. **Boundedness of the normalized polynomials for some weight functions.** Let $P_n[x; f(x)]$ be the polynomial of the nth degree in the normalized orthogonal system on the interval $(-1, 1)$ corresponding to the weight function $f(x)$.

THEOREM I. *If $\rho(x, y) = \rho(y, x) = \rho(-x, -y)$ for all points (x, y) of C, and if the polynomials $P_n[x; \rho(x, 1)]$ are uniformly bounded for all values of n on every closed interval interior to $(-1, 1)$, then the polynomials $p_{nk}(x, y)$ are uniformly bounded for all values of n and k on every closed point set of C not including one of the corner points.*

To prove the theorem it is first necessary to point out certain facts regarding integration over the contour C. The integral of any function $g(x, y)$ over C is given by the formula

$$\int_C g(x, y)ds = \int_{-1}^{1} g(x, 1)dx + \int_{-1}^{1} g(x, -1)dx + \int_{-1}^{1} g(1, y)dy + \int_{-1}^{1} g(-1, y)dy.$$

Hence, if $g(x, y) = -g(y, x)$ at all points of C or if $g(x, y) = -g(-x, -y)$ at all points of C, $\int_C g(x, y)ds = 0$. From this it is possible to prove that each polynomial $p_{nk}(x, y)$ is either symmetric or skew-symmetric and that each is either even or odd; that is, $p_{nk}(x, y) \equiv \pm p_{nk}(y, x)$ and $p_{nk}(x, y) \equiv \pm p_{nk}(-x, -y)$.

The polynomial $p_{2n,1}(x, y)$ is therefore a linear combination of the terms $x^{2n-1}y + xy^{2n-1}, x^{2n-2} + y^{2n-2}, x^{2n-3}y + xy^{2n-3}, \cdots, x^2 + y^2, xy, 1$. It has the following properties of orthogonality:

$$\int_C \rho(x, y)p_{2n,1}(x, y)(x^{2k} + y^{2k})ds = 0, \quad k = 0, 1, \cdots, n-1,$$

$$\int_C \rho(x, y)p_{2n,1}(x, y)(x^{2k-1}y + xy^{2k-1})ds = 0, \quad k = 1, 2, \cdots, n-1.$$

The integrands in these two equations are even and symmetric and the integral over C of such a function is equal to four times the integral of that function over any one side of C. Hence,

$$\int_{-1}^{1} \rho(x, 1)p_{2n,1}(x, 1)(x^{2k} + 1)dx = 0, \qquad k = 0, 1, \cdots, n - 1,$$

$$\int_{-1}^{1} \rho(x, 1)p_{2n,1}(x, 1)(x^{2k-1} + x)dx = 0, \qquad k = 1, 2, \cdots, n - 1.$$

That is, $p_{2n,1}(x, 1)$ is a polynomial in x of degree $2n-1$ which is orthogonal over $(-1, 1)$ with $\rho(x, 1)$ as weight function to $1, x, x^2, \cdots, x^{2n-2}$. Hence,

$$p_{2n,1}(x, 1) = (1/2)P_{2n-1}[x; \rho(x, 1)].$$

By similar methods it can be shown that $p_{2n,2}(x, 1)/(x^2-1)$ is a polynomial of degree $2n-3$ which is orthogonal over $(-1, 1)$ with $(1-x^2)^2\rho(x, 1)$ as weight function to $1, x, x^2, \cdots, x^{2n-4}$; and so

$$p_{2n,2}(x, 1) = (1/2)(x^2 - 1)P_{2n-3}[x; (1 - x^2)^2\rho(x, 1)].$$

The corresponding formulas for the other two polynomials of even degree and the four of odd degree are

$$\begin{aligned}
p_{2n,3}(x, 1) &= (1/2)P_{2n}[x; \rho(x, 1)], \\
p_{2n,4}(x, 1) &= (1/2)(x^2 - 1)P_{2n-2}[x; (1 - x^2)^2\rho(x, 1)], \\
p_{2n+1,1}(x, 1) &= (1/2)(x + 1)P_{2n-1}[x; (1 + x)^2\rho(x, 1)], \\
p_{2n+1,2}(x, 1) &= (1/2)(x - 1)P_{2n-1}[x; (1 - x)^2\rho(x, 1)], \\
p_{2n+1,3}(x, 1) &= (1/2)(x + 1)P_{2n}[x; (1 + x)^2\rho(x, 1)], \\
p_{2n+1,4}(x, 1) &= (1/2)(x - 1)P_{2n}[x; (1 - x)^2\rho(x, 1)].
\end{aligned}$$

It is well known that if the polynomials $P_n[x; f(x)]$ are uniformly bounded on every closed interval interior to $(-1, 1)$, then the polynomials $P_n[x; \pi(x)f(x)]$ are likewise bounded provided $\pi(x)$ is a polynomial which does not vanish on the interior of the interval $(-1, 1)$. Hence, the above formulas, together with the fact that each polynomial $p_{nk}(x, y)$ is symmetric or skew-symmetric and even or odd, establish the theorem.

It is of interest to note that if $\rho(x, y) = |y+x|^\alpha |y-x|^\beta$ where $\alpha > -1$, $\beta > -1$, then the polynomials $p_{nk}(x, 1)$ are expressible in terms of Jacobi polynomials.

3. **Convergence theorems.** Let $f(x, y)$ be any function of x and y

which is integrable in the sense of Lebesgue over C and let us consider the series

$$(1) \qquad \sum_{n=0}^{\infty} \sum_{k=1}^{4} a_{nk} p_{nk}(x, y)$$

where $a_{nk} = \int_C \rho(u, v) f(u, v) p_{nk}(u, v) ds$ and where (x, y) is any point of C. (For convenience of notation we define p_{02}, p_{03}, p_{04}, p_{13}, p_{14}, and p_{24} to be identically zero.) Let $K_n(x, y, u, v) = \sum_{m=0}^{n} \sum_{k=1}^{4} p_{mk}(x, y) p_{mk}(u, v)$, and let $S_n(x, y)$ be the partial sum of (1) through terms of the nth degree. Then

$$(2) \quad S_n(x, y) - f(x, y) = \int_C \rho(u, v) [f(u, v) - f(x, y)] K_n(x, y, u, v) ds.$$

(In formulas where x, y, u, and v appear under the integral sign u and v are to be taken as the variables of integration.) By using a formula[3] for $K_n(x, y, u, v)$ analogous to the Christoffel-Darboux identity, the difference $S_n(x, y) - f(x, y)$ can be expressed as a sum of at most thirty-two terms of the form

$$(3) \qquad k p_{jh}(x, y) \int_C \rho(u, v) \frac{f(u, v) - f(x, y)}{(Au + Bv) - (Ax + By)} p_{mk}(u, v) ds$$

where A and B are any arbitrary constants, where k is a constant numerically less that $|A| + |B|$, and where m and j take on the values n and $n+1$. The problem of convergence at a specified point thus reduces to proving that an expression of the form (3) approaches zero as n becomes infinite.

LEMMA 1. *If the weight function $\rho(x, y)$ is of such a nature that the corresponding polynomials are uniformly bounded on every closed point set of C not containing a corner point and if $\phi(x, y)$ is a function such that $\rho\phi$ is integrable over C and $\rho\phi^2$ is integrable over some neighborhood on C of each corner point, then*

$$\lim_{n \to \infty} \int_C \rho(u, v) \phi(u, v) p_{nk}(u, v) ds = 0, \quad k = 1, 2, 3, 4.$$

The proof of this lemma is omitted since it can be easily constructed from proofs already known.[4] On the basis of this lemma we can now prove the following assertion:

[3] D. Jackson, (1), p. 433.

[4] See, for example, D. Jackson, *Series of orthogonal polynomials*, Annals of Mathematics, (2), vol. 34 (1933), pp. 536–537.

THEOREM II. *Assume*

(a) *the weight function $\rho(x, y)$ is such that the corresponding polynomials $p_{nk}(x, y)$ are uniformly bounded on every closed point set of C not containing a corner point*;

(b) *(x, y) is a point of C, not a corner*;

(c) *the function ρf is integrable over C and ρf^2 is integrable over a neighborhood on C of each corner point*;

(d) *the function*

$$\rho(u, v)\frac{f(u, v) - f(x, y)}{(u - x) + (v - y)}$$

is integrable over some neighborhood on C of the point (x, y).

Then the series (1) *converges at the point (x, y) to the value $f(x, y)$.*

Let us first make the additional hypothesis that there is a point (x', y') on C on the interior of a side different from the one containing (x, y) such that the function

$$\rho(u, v)\frac{f(u, v) - f(x, y)}{(u - x') + (v - y')}$$

is integrable over some neighborhood on C of the point (x', y'). Determine A and B so that the straight line $A(u-x)+B(v-y)=0$ in the uv plane passes through the point (x', y'). Then in some neighborhood on C of (x, y) the expression $A(u-x)+B(v-y)$ is a non-vanishing constant multiple of $u-x+v-y$ and in some neighborhood of (x', y'), of $u-x'+v-y'$. If we let

$$\phi(u, v) = \frac{f(u, v) - f(x, y)}{A(u - x) + B(v - y)},$$

then $\rho\phi$ is integrable over C and $\rho\phi^2$ is integrable over a neighborhood of each corner point; hence, by Lemma 1, the entire expression (3) approaches zero as n becomes infinite, which insures the convergence of (1) to the value $f(x, y)$.

To prove the theorem in the general case let us choose two distinct points (x', y') and (x'', y''), each on the interior of a side of C different from the side containing (x, y), and let us define a function $g(u, v)$ to be equal to $f(u, v)$ throughout neighborhoods of (x, y) and (x', y'), equal to $g(x, y)=f(x, y)$ throughout a neighborhood of (x'', y''), and equal to zero elsewhere on C. Then the series of form (1) corresponding to the functions $f-g$ and g will converge at the point (x, y) to the values $f(x, y)-g(x, y)$ and $g(x, y)$ respectively. Hence, the series for f will converge at the point (x, y) to the value $f(x, y)$.

COROLLARY. *If the function f is such that ρf is integrable over C and ρf^2 is integrable over a neighborhood of each corner point, the behavior of the series* (1) *at a point (x, y) on the interior of a side of C depends only on the behavior of the function $f(u, v)$ in the neighborhood of that point.*

A study of the uniform convergence of (1) to a continuous function $f(x, y)$ can be made with the same methods that have been used in a study of Legendre series.[5] In applying these methods it is first necessary to consider the degree of polynomial approximation obtainable over the contour C to a function which is continuous on C. By adaptation of existing theorems[6] with regard to polynomial approximation on an interval, the following conclusion is obtained:

THEOREM III. *If the polynomials $p_{nk}(x, y)$ are uniformly bounded and the weight function $\rho(x, y)$ bounded on every closed point set of C not including one of the corner points and if $f(x, y)$ is a function such that $f(x, \pm 1)$ and $f(\pm 1, x)$ have continuous pth derivatives $(p \geqq 0)$ on the interval $(-1, 1)$ with moduli of continuity at most equal to $\omega(\delta)$; then, for every $\eta > 0$,*

$$|S_n(x, y) - f(x, y)| \leqq \frac{K \log n}{n^p} \omega(1/n)$$

for all (x, y) on C not within a distance η of any corner point and for all $n > 1$, the constant K being independent of n and of (x, y) but dependent on η.

4. **Other contours.** The methods used in the convergence proofs of the preceding section are applicable to orthogonal systems over a highly general contour provided the normalized polynomials of the system possess the requisite property of boundedness on that contour. There has been as yet, however, no general method devised for establishing the latter property. Two other ranges of integration for which the result can be established by reasoning similar to that already used are a pair of bisecting line segments and a pair of concentric circles.

An orthogonal system over the first of these domains will have two polynomials of the nth degree for $n \geqq 1$. If the range is taken as consisting of the portions of the x and y axes contained in the intervals $-1 \leqq x \leqq 1$, $-1 \leqq y \leqq 1$, the two polynomials of nth degree of a normalized, orthogonal system corresponding to a symmetric weight

[5] See, for example, D. Jackson, *The Theory of Approximation*, American Mathematical Society Colloquium Publications, vol. 11, New York, 1930, pp. 25–32.

[6] D. Jackson, loc. cit., pp. 13–18.

function $\rho(x, y)$ are given by the formulas

$$p_{n1}(x, 0) = 2^{-1/2}P_n[x; \rho(x, 0)], \qquad p_{n2}(x, 0) = 2^{-1/2}xP_{n-1}[x; x^2\rho(x, 0)],$$
$$p_{nk}(x, y) = (-1)^{k-1}p_{nk}(y, x).$$

The convergence theorems for this system are similar to those of the last section with the points $(\pm 1, 0), (0, \pm 1), (0, 0)$ appearing as exceptional points analogous to the corner points of the square.

The contour made up of two concentric circles is an example of a contour which has no singular points. An orthogonal system over this contour will have four polynomials of the nth degree for $n \geqq 3$. If the two circles are taken with center at the origin and radii a and b, the polynomials of a normalized system corresponding to the weight function 1 are represented by the formulas

$$p_{nk}(a \cos \theta, a \sin \theta) = \alpha_{nk1} \cos n\theta + \alpha_{nk2} \cos (n - 2)\theta,$$
$$p_{nk}(b \cos \theta, b \sin \theta) = \beta_{nk1} \cos n\theta + \beta_{nk2} \cos (n - 2)\theta, \quad k = 1, 3.$$

The formulas for p_{n2} and p_{n4} are of the same form with the cosines replaced by sines. The constants α_{nkm} and β_{nkm} involved in the formulas are bounded for all values of n, k, and m; so the polynomials themselves are uniformly bounded over the entire contour. The convergence theorems for this system of polynomials can thus be formulated without reference to any exceptional points.

THE UNIVERSITY OF MINNESOTA

ON THE CONVERSE OF THE TRANSITIVITY OF MODULARITY

Y. K. WONG

E. H. Moore's theorem on the transitivity of modularity is as follows: Consider the basis[1] $\mathfrak{A}$, $\mathfrak{P}$, ϵ; if a positive hermitian matrix ϵ_0 is modular as to ϵ ϵ, then every vector which is modular as to ϵ_0 is modular[2] as to ϵ (that is, $\mathfrak{M}_{\epsilon_0} \subset \mathfrak{M}_\epsilon$).

In his doctoral thesis, the author establishes the converse of the preceding theorem as a consequence of the Hellinger-Toeplitz theorem.[3] In this note, we give a new proof for the converse of the transitivity of modularity, and then deduce the generalized Hellinger-Toeplitz theorem as a corollary. The converse of the transitivity of modularity is, therefore, equivalent to the Hellinger-Toeplitz theorem. We also establish the converse of the transitivity of modularity for matrices, and a theorem on the transitivity of accordance and finiteness.

THEOREM I. *Consider the basis* $\mathfrak{A}$, $\mathfrak{P}$, ϵ; *and let* ϵ_0 *be a positive hermitian matrix. Then the following assertions are equivalent*:

(1) *every vector* μ_0 *modular as to* ϵ_0 *is modular as to* ϵ;

(2) ϵ_0 *is modular as to* ϵ ϵ_0;

(3) ϵ_0 *is modular as to* ϵ ϵ.

If one of the preceding conditions is satisfied, the modulus of ϵ_0 *as to* ϵ ϵ *is equal to the norm of* ϵ_0 *as to* ϵ ϵ_0.

In the course of demonstration, we let $\mathfrak{M}_0$ denote the space of vectors μ_0 modular as to ϵ_0; J_0, the integration process based on ϵ_0; and M_0, the modulus as to ϵ_0. Similar interpretations are given to the symbols $\mathfrak{M}$, J, M, for the base matrix ϵ. A vector which is finite as to ϵ is denoted by β.

If every μ_0 is modular as to ϵ, the matrix ϵ_0 is of type $\mathfrak{M}_0\overline{\mathfrak{M}}$. Then $J\epsilon_0\beta$ is in $\mathfrak{M}_0$ for every β, and $J_0(J\bar{\beta}\epsilon_0)\mu_0 = J\bar{\beta}J_0\epsilon_0\mu_0 = J\bar{\beta}\mu_0$ for every pair β, μ_0. Consequently, for every β, $M_0J\epsilon_0\beta$ is equal to the least upper bound of $|J\bar{\beta}\mu_0|$ for all μ_0 such that $M_0\mu_0 \leqq 1$, by part (2) of Theorem (41.9) in G.A. Similarly, for every μ_0, which is modular as to ϵ by hypothesis, $M\mu_0$ is equal to the least upper bound of $|J\bar{\beta}\mu_0|$

[1] E. H. Moore, *General Analysis* (G.A. for abbreviation), Part I, p. 4, and Part II, p. 84.

[2] Theorem (46.4), part (1) in G.A., II, p. 137.

[3] *Spaces associated with non-modular matrices with applications to reciprocals*, Chicago thesis, 1931, pp. 3–9. The same proof is given in G.A., II, p. 193.

for all β such that $M\beta \leqq 1$. If the class $\mathfrak{L}$ is identified as the class of vectors $\bar{\beta}$ such that $M\beta \leqq 1$, and F_l, on $\mathfrak{M}_0$ to $\mathfrak{N}$, is defined to be $(|J\bar{\beta}\mu_0|\,|\mu_0)$ for every $l=\bar{\beta}$, then by Theorem (53.55) in G.A.,[4] the upper bound of $M_0 J\epsilon_0\beta$, for all β finite as to ϵ such that $M\beta \leqq 1$, is finite. By Theorem (46.85) in G.A., ϵ_0 is modular as to ϵ_0 ϵ. Since ϵ_0 is hermitian, ϵ_0 is also modular[5] as to ϵ ϵ_0.

When condition (2) is true, then condition (3) is secured by a simple application[6] of the composition of modularity[7] to $\epsilon_0 = J_0\epsilon_0\epsilon_0$. That the last condition implies the first is proved in Theorem (46.4) of Moore's G.A.

From $\epsilon_0 = J_0\epsilon_0\epsilon_0$ and part (2) of Theorem (46.9) in G.A., we have $N_{\epsilon_0\epsilon}\epsilon_0 = M_{\epsilon\epsilon}J_0\epsilon_0\epsilon_0 = M_{\epsilon\epsilon}\epsilon_0$. This completes the proof.

The hypothesis of the preceding theorem is assumed for the following corollary:

COROLLARY. *Let $\mathfrak{M}_{0*}$ consist of all μ_0 whose moduli as to ϵ_0 are bounded by a fixed constant. If $\mathfrak{M}_{0*}$ is a subset of $\mathfrak{M}$, then the moduli as to ϵ of all vectors in $\mathfrak{M}_{0*}$ are also bounded.*

We may assume, without losing generality, that the moduli of all vectors in $\mathfrak{M}_{0*}$ are at most unity. Since the spaces $\mathfrak{M}_0$ and $\mathfrak{M}$ are linear, the condition that every μ_0 for which $M_0\mu_0 \leqq 1$ is modular as to ϵ is equivalent to condition (1) in the preceding theorem. Consequently ϵ_0 is modular as to ϵ ϵ_0. The equation $\mu_0 = J_0\epsilon_0\mu_0$ gives, by Theorem (46.7) in G.A., that $M\mu_0 \leqq M_{\epsilon_0\epsilon_0}\epsilon_0$ whenever $M_0\mu_0 \leqq 1$.

THEOREM II. *Consider the basis $\mathfrak{A}$, $\mathfrak{P}^1$, $\mathfrak{P}^2$, ϵ^1, ϵ^2; and let ϵ_0^1, ϵ_0^2 be two positive hermitian matrices. Then the following assertions are equivalent:*

(1) *every matrix κ^{12} modular as to ϵ_0^1 ϵ_0^2 is of type $\mathfrak{M}^1\overline{\mathfrak{M}}^2$;*

(2) *ϵ_0^1 is modular as to ϵ^1 ϵ^1, and ϵ_0^2 is modular as to ϵ^2 ϵ^2;*

(3) *every matrix κ^{12} modular as to ϵ_0^1 ϵ_0^2 is modular as to ϵ^1 ϵ^2.*

For the demonstration of the theorem, we shall show that (1)→(2)→(3)→(1). The second implication is proved in part (2) of Theorem (46.4) in G.A. The last implication follows from the fact that every matrix κ^{12} modular as to ϵ^1 ϵ^2 is of type $\mathfrak{M}^1\overline{\mathfrak{M}}^2$. To show

[4] See also Hildebrandt, *On uniform limitedness of sets of functional operations*, this Bulletin, vol. 29 (1923), pp. 309–315; Fréchet, *Sur les fonctionelles bilinéaires*, Transactions of this Society, vol. 16 (1915), pp. 217–218.

[5] By a similar reasoning, we may, of course, deduce the Hellinger-Toeplitz theorem as a consequence of Theorem (53.55).

[6] See the author's thesis, loc. cit., p. 8, or Moore, G.A., II, p. 193.

[7] Moore, G.A., II, p. 144.

the first implication, consider any μ_0^1, $\mu_0^2 \neq 0^2$ which are modular as to ϵ_0^1, ϵ_0^2 respectively, Theorem (47.2) in G.A. shows that $\mu_0^1\bar{\mu}_0^2$ is modular as to ϵ^1 ϵ^2, and hence by hypothesis, $\mu_0^1\bar{\mu}_0^2$ is of type $\mathfrak{M}^1\overline{\mathfrak{M}}^2$. Since $\mu^2 \neq 0^2$, let $a \equiv \mu^2(p^2) \neq 0$. Then $\mu_0^1 \cdot a$, and hence μ_0^1, is modular as to ϵ^1. This proves that every μ_0^1 modular as to ϵ_0^1 is modular as to ϵ^1. By Theorem I, ϵ_0^1 is modular as to ϵ^1 ϵ^1. Similarly, we prove that ϵ_0^2 is modular as to ϵ^2 ϵ^2. The proof was suggested by Dr. Coral.

THEOREM III. (Generalized Hellinger-Toeplitz theorem.) *Consider the basis* $\mathfrak{A}$, $\mathfrak{P}^1$, $\mathfrak{P}^2$, ϵ^1, ϵ^2. *A matrix* κ^{12} *is modular as to* ϵ^1 ϵ^2 *if and only if* κ^{12} *is by rows of* $\overline{\mathfrak{M}}^2$ *and* $J^2\kappa^{12}\mu^2$ *is modular as to* ϵ^1 *for every* μ^2.

To prove the theorem, we make use of the fact that κ^{12} is modular as to ϵ^1 ϵ^2 if and only if the following condition holds:

(M) κ^{12} *is by rows of* $\overline{\mathfrak{M}}^2$, *and* $J^2\kappa^{12}\kappa^{*21}$ *is modular as to* ϵ^1 ϵ^1.

This is Theorem (46.9) in G.A., with the omission of the redundant condition that κ^{12} is by columns accordant as to ϵ^1. (For when κ^{12} satisfies the conditions (M), κ^{12} is by columns A^1. To prove this, we note that $J^2\kappa^{12}\kappa^{*21}$ is A^{11} by Theorem (46.65) in G.A. Consequently, when $S_\sigma^1\epsilon^1\alpha^1 = 0^1$, then $J^2(S_\sigma^1\bar{\alpha}^1\kappa^{12}, S_\sigma^1\kappa^{*21}\alpha^1) = S_\sigma^1 S_\sigma^1\bar{\alpha}^1 J^2\kappa^{12}\kappa^{*21}\alpha^1 = 0$, which implies that $S_\sigma^1\bar{\alpha}^1\kappa^{12} = 0^2$, since J^2 is proper. Thus κ^{12} is by columns A^1.) Consequently, it suffices to prove the following statement: *When* κ^{12} *is by rows of* $\overline{\mathfrak{M}}^2$, *the matrix* $J^2\kappa^{12}\kappa^{*21}$ *is modular as to* ϵ^1 ϵ^1 *if and only if* $J^2\kappa^{21}\mu^2$ *is modular as to* ϵ^1 *for every* μ^2.

Using the notation introduced by E. H. Moore in his study of generalized Fourier theory, we denote the positive hermitian matrix $J^2\kappa^{21}\kappa^{*21}$ by ϵ_*^1. It was shown by Moore that the space of vectors modular as to ϵ_*^1 is equal[8] to the space of vectors $J^2\kappa^{12}\mu^2$ for all μ^2 in $\mathfrak{M}^2$. When κ^{12} is assumed to be by rows of $\overline{\mathfrak{M}}^2$, the assertion that $J^2\kappa^{12}\mu^2$ is modular as to ϵ^1 for every μ^2 is equivalent to the assertion that every vector modular as to ϵ_*^1 is modular as to ϵ^1. By Theorem I, the latter assertion is valid if and only if ϵ_*^1 is modular as to ϵ^1 ϵ^1. This proves the theorem.

The basis stated in the preceding theorem is assumed for the following corollary:

COROLLARY. *Suppose that* κ^{12} *is by rows of* $\overline{\mathfrak{M}}^2$. *Then* κ^{12} *is modular as to* ϵ^1 ϵ^2 *if and only if every vector modular as to* $J^2\kappa^{12}\kappa^{*21}$ *is modular as to* ϵ^1.

The transitivity for accordance and finiteness is stated in the following theorem:

[8] Moore, G. A., I, p. 22.

THEOREM IV. *Consider the basis* $\mathfrak{A}$, $\mathfrak{B}$, ϵ *and let* ϵ_0 *be a positive hermitian matrix. Then*

(a) *every vector accordant as to* ϵ_0 *is accordant as to* ϵ *if and only if* ϵ_0 *is accordant as to* $\epsilon\,\epsilon$;

(b) *every vector finite as to* ϵ_0 *is finite as to* ϵ *if and only if* ϵ_0 *is of type* $F\overline{F}$.

In part (a), if every vector accordant as to ϵ_0 is accordant as to ϵ, then ϵ_0, being of type $A_0\overline{A}_0$, is of type $A\overline{A}$. By Theorem (46.5) in G.A., ϵ_0 is accordant as to $\epsilon\,\epsilon$. Conversely every vector ξ accordant as to ϵ_0 satisfies the relation $\xi = J_0\epsilon_0\xi = L_\sigma J_0\epsilon_0\xi_\sigma$. Now $J_0\epsilon_0\xi_\sigma$, being a finite (right) linear combination of the columns of ϵ_0, is a vector accordant as to ϵ for every σ. By Theorem (48.2) in G.A., ξ is accordant as to ϵ. Part (b) is an immediate consequence of the definition of finiteness.

THE UNIVERSITY OF CHICAGO

ON THE EXTENSION OF A FUNCTIONAL INEQUALITY OF S. BERNSTEIN TO NON-ANALYTIC FUNCTIONS[1]

R. J. DUFFIN AND A. C. SCHAEFFER

We wish to demonstrate here the following elementary inequality of the differential calculus. If a function[2] satisfies the conditions $\{f(x)\}^2 \leqq 1$ and $\{f^{(n)}(x)\}^2 + \{f^{(n-1)}(x)\}^2 \leqq 1$ for all x and for some positive integer n, then the latter inequality is valid also when n is replaced by any smaller positive integer.

That such an inequality might be true is suggested by the validity of a similar but more specialized inequality concerning trigonometric polynomials.[3] Thus, if $P(x) = \sum_0^N \{a_\nu \cos(\nu x/N) + b_\nu \sin(\nu x/N)\}$ and if $\{P(x)\}^2 \leqq 1$ for all x, it has been proved that $\{P^{(k)}(x)\}^2 + \{P^{(k-1)}(x)\}^2 \leqq 1$, $(k = 1, 2, 3, \cdots)$. This theorem, a refinement of a theorem of S. Bernstein, has been proved by several different methods,[4] and generalizations have been given which prove that the inequality is true for a wider class of analytic functions. It will be shown here that this theorem is a rapid deduction from the elementary inequality given above. Moreover, this method of proof serves to distinguish those features of Bernstein's theorem arising from the characteristic properties of trigonometric polynomials from those which are merely properties of the differential coefficient.

The second part of this paper is concerned with finding the functions which cause the inequality to become an equality at some point. For example, if $\{f^{(3)}(x)\}^2 + \{f^{(2)}(x)\}^2 \leqq 1$ and $\{f(x)\}^2 \leqq 1$, we find that the equality $f'(0) = 1$ necessitates that $f(x) \equiv \sin x$. It is to be noted

[1] Presented to the Society, September 6, 1938, under the title *Extension of a theorem of S. Bernstein to non-analytic functions.*

[2] In stating that an inequality involving $f^{(n)}(x)$ is satisfied in an interval we imply that the derivatives of all orders up to and including the nth exist in the same interval. All functions and constants appearing in this paper are considered real unless the contrary is explicitly stated.

[3] In formulating such a comparison, we have taken a hint from a paper by Oystein Ore, *On functions with bounded derivatives*, Transactions of this Society, vol. 43 (1938), pp. 321–326. Ore introduces an "extension principle" which indicates how certain theorems concerning ordinary algebraic polynomials may be modified in order to be applicable to arbitrary functions.

[4] G. Szegö, Schriften der Königsberger gelehrten Gesellschaft, 5th year, no. 4 (1928), pp. 59–70. J. van der Corput and G. Schaake, Compositio Mathematica, vol. 2 (1935), pp. 321–361; vol. 3 (1936), p. 128. R. Boas, Transactions of this Society, vol. 40 (1936), pp. 287–308. R. Duffin and A. Schaeffer, this Bulletin, vol. 43 (1937), pp. 554–556; vol. 44 (1938), pp. 236–240.

that the behavior of this system of differential inequalities is similar to the behavior of a differential equation because the assignment of certain initial values gives rise to a unique solution. It is possible to construct other examples of differential inequalities with this stability property.

The method of proof employs only fundamental notions. Essentially all that is needed is the mean value theorem and Weierstrass' theorem on the maximum of a continuous function.

THEOREM A. *If $f(x)$ is a function which for some integer n and for all x satisfies the conditions*

(a) $$\{f(x)\}^2 \leqq 1,$$

(b) $$\{f^{(n)}(x)\}^2 + \{f^{(n-1)}(x)\}^2 \leqq 1,$$

then for $k=1, 2, 3, \cdots, n$ and for all x

(c) $$\{f^{(k)}(x)\}^2 + \{f^{(k-1)}(x)\}^2 \leqq 1.$$

PROOF. Considering first the case $n=2$, we have by hypothesis $\{f(x)\}^2 \leqq 1$ and $\{f'(x)\}^2 + \{f''(x)\}^2 \leqq 1$. It is to be shown that the function $\phi(x) \equiv \{f(x)\}^2 + \{f'(x)\}^2$ is not greater than 1. Suppose, on the contrary, that at some point x_1, $\phi(x_1) = 1+2\delta$, $\delta > 0$. If x is any point to the right of x_1, then by the mean value theorem a point x_2 exists such that

$$f'(x_2) = \{f(x) - f(x_1)\}/(x - x_1), \qquad x_1 < x_2 < x.$$

Taking $x - x_1 = 2/(\delta)^{1/2}$, we see that $|f'(x_2)| \leqq (\delta)^{1/2}$. Hence $\phi(x_2) \leqq 1 + \{f'(x_2)\}^2 \leqq 1+\delta$. The same method is available to show that there is a point to the left of x_1 where $\phi(x) \leqq 1+\delta$. Thus, $\phi(x)$ is greater at x_1 than it is at certain points to the left and to the right; it must, therefore, have a maximum in the interior of the interval defined by these points. At this maximum point, say x_0, we have $\phi(x_0) \geqq 1+2\delta$; and the derivative is equal to zero:

$$2f'(x_0)\{f(x_0) + f''(x_0)\} = 0.$$

But, if the first factor is zero, then $\phi(x_0) = \{f(x_0)\}^2 \leqq 1$, while if the second factor is zero, then

$$\phi(x_0) = \{-f''(x_0)\}^2 + \{f'(x_0)\}^2 \leqq 1.$$

This contradiction proves the theorem for $n=2$.

For n greater than 2 we proceed by induction. Thus, suppose the theorem is true for $n=m-1$ and consider the case $n=m$. First, in order to show that $f'(x)$ is bounded uniformly in x, we obtain the

crude estimate $|f'(x_0)| \leqq 5$ for an arbitrary point x_0. By the mean value theorem

$$f(x_0 + 1) - f(x_0) = f'(x_0 + h), \qquad 0 < h < 1.$$

Because f is bounded by 1, this gives, for some h between 0 and 1,

$$|f'(x_0 + h)| \leqq 2. \tag{1}$$

Again, by the mean value theorem, if x is any point and h is the same constant appearing in (1), then for some point x_1

$$f^{(m-2)}(x + h) - f^{(m-2)}(x) = hf^{(m-1)}(x_1). \tag{2}$$

But, by hypothesis $f^{(m-1)}$ is bounded by 1; so, using (2), we obtain

$$\{f^{(m-2)}(x + h) - f^{(m-2)}(x)\}^2 + \{f^{(m-1)}(x + h) - f^{(m-1)}(x)\}^2 \leqq 9.$$

Note that $\{f(x+h)-f(x)\}^2<9$ follows from hypothesis (a), so the theorem is applicable to the function $\{f(x+h)-f(x)\}/3$ in the case $n=m-1$ for which it has been supposed true. As a consequence

$$\left\{\frac{f'(x + h) - f'(x)}{3}\right\}^2 + \left\{\frac{f(x + h) - f(x)}{3}\right\}^2 \leqq 1.$$

In particular the last relation gives for the point x_0 the inequality

$$|f'(x_0 + h) - f'(x_0)| \leqq 3. \tag{3}$$

Combining (1) and (3) we obtain the desired bound $|f'(x_0)| \leqq 5$.

Let β be the least upper bound of $|f'(x)|$ for all values of x; we have just shown that β is finite. If $\beta>1$, the function $\beta^{-1}f'(x)$ satisfies the conditions of the theorem in the case $n=m-1$ and consequently

$$\{\beta^{-1}f''(x)\}^2 + \{\beta^{-1}f'(x)\}^2 \leqq 1. \tag{4}$$

Let $F(x)=\beta^{-1}f(x)+\gamma(1-\beta^{-1})$ where $\gamma=+1$ or -1 and is to be determined later. Because $\{f(x)\}^2\leqq 1$, clearly $\{F(x)\}^2\leqq 1$; and by (4), $\{F''(x)\}^2+\{F'(x)\}^2\leqq 1$. Applying the theorem to the function $F(x)$ in the case $n=2$, we have $\{F'(x)\}^2+\{F(x)\}^2\leqq 1$ or

$$\{\beta^{-1}f'(x)\}^2 + \{\beta^{-1}f(x)\}^2 + 2\gamma\beta^{-1}(1 - \beta^{-1})f(x) + (1 - \beta^{-1})^2 \leqq 1.$$

By the definition of β there exist points for which the first term on the left is arbitrarily close to 1. And for fixed x the term containing γ can be made nonnegative since γ is either $+1$ or -1. Thus, the least upper bound of the left-hand side is not smaller than $1+(1-\beta^{-1})^2$; and because this can be no larger than 1, it follows that $\beta\leqq 1$.

The problem is by the above considerations reduced to the case $\{f'(x)\}^2 \leqq 1$. Applying the theorem to the function $f'(x)$ with $n=m-1$, we have for $k=2, 3, 4, \cdots, m-1, m$

$$\{f^{(k)}(x)\}^2 + \{f^{(k-1)}(x)\}^2 \leqq 1. \tag{c}$$

But, now we may apply the theorem to $f(x)$ with $n=2$, and this shows that (c) is true for $k=1$ also.[5]

We wish now to determine the class of functions which make relation (c) of Theorem A an equality at some point and for some value of k. The following lemma is of central importance for this investigation.

LEMMA. *If the function $f(x)$ satisfies in the interval $0 \leqq \alpha \leqq x \leqq \pi/2$ the conditions*

$$\{f''(x)\}^2 + \{f'(x)\}^2 \leqq 1, \tag{a}$$

$$f(\alpha) = \sin \alpha, \qquad f'(\alpha) = \cos \alpha, \tag{b}$$

$$f(\pi/2) \leqq 1, \tag{c}$$

then in the same interval

$$f(x) \equiv \sin x. \tag{d}$$

PROOF. It is convenient to introduce two auxiliary functions $\omega(x)$ and $\psi(x)$, defined by the relations

$$\omega(x) = f(x) - \sin x,$$

$$2\psi(x) = 1 + \{\omega''(x)\}^2 + \{\omega'(x)\}^2 - \{f''(x)\}^2 - \{f'(x)\}^2.$$

By condition (b) $\omega(\alpha)=\omega'(\alpha)=0$, and by (a) $\psi(x)$ is nonnegative in the interval under consideration. Writing $f(x)=\omega(x)+\sin x$ and substituting in the expression for $2\psi(x)$ gives

$$\psi(x) = \sin x\, \omega''(x) - \cos x\, \omega'(x);$$

and if $x \neq 0$, this relation may be put in the form

$$\frac{d}{dx}(\omega'(x)/\sin x) = \psi(x)/\sin^2 x. \tag{1}$$

[5] The following three modifications of Theorem A are of some interest: (1) The assumption that $f^{(n)}(x)$ exists may be replaced by weaker conditions. For example, in condition (b), $\{f^{(n)}(x)\}^2$ may be interpreted as $\limsup_{h \to \pm 0}\{(f^{(n-1)}(x+h)-f^{(n-1)}(x))/h\}^2$. (2) The theorem may be stated for a finite interval provided the inequalities (c) are given to be satisfied at the end points of the interval for every value of k. (3) The theorem is valid for complex functions $f(x)$ provided $|f(x)| \leqq 1$ and relations of the form $\{f^{(R)}(x)\}^2+\{f^{(R-1)}(x)\}^2 \leqq 1$ are replaced by relations of the form $|f^{(R)}(x)\cos\alpha + f^{(R-1)}(x)\sin\alpha| \leqq 1$ for all real x and α.

Considering first the case $\alpha>0$, we have by the mean value theorem if $0<\alpha<x<\pi/2$

$$\omega'(x)/\sin x - \omega'(\alpha)/\sin \alpha = (x-\alpha)\psi(x_1)/\sin^2 x_1, \qquad \alpha < x_1 < x.$$

But, $\omega'(\alpha)=0$ and $\psi(x)\geqq 0$; so

$$\omega'(x) \geqq 0, \qquad \alpha \leqq x \leqq \pi/2, \tag{2}$$

which implies, by the mean value theorem, that $\omega(x)$ is non-decreasing in this interval. Because ω vanishes at $x=\alpha$ and because, by (c), ω cannot be greater than zero at $x=\pi/2$, it follows from (2) that $\omega(x)$ vanishes identically in the interval $(\alpha, \pi/2)$.

Special attention must be given the case $\alpha=0$. If $x_0<x$ are any two points in the interior of the interval $(0, \pi/2)$, by the mean value theorem the relation (1) gives

$$\omega'(x)/\sin x - \omega'(x_0)/\sin x_0 = (x-x_0)\psi(x_1)/\sin^2 x_1,$$
$$0 < x_0 < x_1 < x < \pi/2.$$

Consider the limiting form of this relation as x_0 approaches zero for fixed x. The second term on the left approaches $\omega''(0)$, and the expression on the right, being always nonnegative, must have a nonnegative limit. By conditions (a) and (b) it is clear that $f''(\alpha)=0$ if $\alpha=0$, and this implies that $\omega''(0)=0$. Thus relation (2) is valid also for the case $\alpha=0$, and $\omega(x)$ must vanish identically in the interval $(0, \pi/2)$; this completes the proof.

THEOREM B. *If $f(x)$ is a function which for all x satisfies the conditions*

$$\{f(x)\}^2 \leqq 1, \tag{a}$$

$$\{f''(x)\}^2 + \{f'(x)\}^2 \leqq 1, \tag{b}$$

and if at some point x_1

$$\{f'(x_1)\}^2 + \{f(x_1)\}^2 = 1, \tag{c}$$

then for some constant γ

$$f(x) \equiv \sin(x+\gamma) \tag{d}$$

in an interval from x_1 to the nearest point where $|\sin(x+\gamma)|=1$, and in case $f(x_1)=0$ the interval extends a distance $\pi/2$ on each side of x_1.

PROOF. There is no loss of generality in supposing $f(x_1)\geqq 0, f'(x_1)\geqq 0$ because the general case may be reduced to this by considering the function $\pm f(\pm x)$ for some choice of signs. (If $f(x_1)=0$, the signs can

be chosen in two ways.) Thus let $f(x_1)=\sin\alpha$ and $f'(x_1)=\cos\alpha$, $0\leqq\alpha\leqq\pi/2$. Again there is no loss of generality if we assume $x_1=\alpha$ since this situation may be realized by a suitable translation of the origin. Application of the preceding lemma completes the proof of Theorem B.

THEOREM C. *If $f(x)$ is a function which for all x and for some integer n greater than 2 satisfies the conditions*

(a) $$\{f(x)\}^2 \leqq 1,$$

(b) $$\{f^{(n)}(x)\}^2 + \{f^{(n-1)}(x)\}^2 \leqq 1,$$

and if at some point x_1 and for some integer k in the range $1\leqq k\leqq n-1$,

(c) $$\{f^{(k)}(x_1)\}^2 + \{f^{(k-1)}(x_1)\}^2 = 1,$$

then there is a constant γ such that for all x

(d) $$f(x) \equiv \sin(x+\gamma)$$

except possibly if $k=1$ and $\{f(x_1)\}^2=1$.

PROOF. We shall tacitly employ Theorem A throughout. First consider the case $k=1$; suppose

$$\{f'(x_1)\}^2 + \{f(x_1)\}^2 = 1, \qquad \{f(x_1)\}^2 < 1.$$

By Theorem B, $f(x)\equiv\sin(x+\gamma)$ throughout an interval at one end of which $f(x)=\pm 1$. At this end point, say x_2, it follows by continuity that $\{f'(x_2)\}^2+\{f(x_2)\}^2=1$ and $f'(x_2)=0$. Applying Theorem B to the function $f'(x)$, we obtain $f'(x)\equiv\cos(x+\gamma)$ over an interval extending a distance $\pi/2$ on each side of x_2. At the end points of this interval $\{f'(x)\}^2+\{f(x)\}^2=1$ and $f(x)=0$; so applying Theorem B to the function $f(x)$, we see that $f(x)\equiv\sin(x+\gamma)$ in an interval of length 2π. Clearly this process may be repeated to give a continuation of $f(x)$ for all values of x, and this proves the Theorem for $k=1$.

The problem for higher values of k is easily reduced to the case $k=1$ just discussed. Suppose for some x_1 and k

(1) $$\{f^{(k)}(x_1)\}^2 + \{f^{(k-1)}(x_1)\}^2 = 1, \qquad 1 < k < n.$$

By application of Theorem B to the function $f^{(k-1)}(x)$ we see that there is a point where $f^{(k-1)}(x)=\pm 1$. At this point we have

(2) $$\{f^{(k-1)}(x)\}^2 + \{f^{(k-2)}(x)\}^2 = 1, \qquad f^{(k-2)}(x) = 0.$$

Now, note that (2) is the same as (1) with k replaced by $k-1$ and with the added condition that $f^{(k-2)}(x)$ is zero. By repeating this proc-

ess a sufficient number of times we shall find that at some point $\{f'(x)\}^2+\{f(x)\}^2=1$ and $f(x)=0$. This is an instance of the case already discussed, and the theorem is proved.

We are now able to give a proof of S. Bernstein's theorem on trigonometric polynomials, which may be stated in the following form:[6]

If $\phi(x)$ is a trigonometric polynomial of degree N or less, $\phi(x)=\sum_0^N(a_\nu\cos\nu x+b_\nu\sin\nu x)$; and if $|\phi(x)|\leqq 1$ for all x, then $|\phi'(x)|<N$ for all x, unless $\phi(x)\equiv\sin(Nx+\alpha)$ for some constant α.

PROOF. Let $f(x)=\phi(x/\lambda)$ where λ is some constant greater than N. Then as $n\to\infty$, $f^{(n)}(x)\to 0$ uniformly in $(-\infty,\infty)$; hence for sufficiently large n, $\{f^{(n)}(x)\}^2+\{f^{(n-1)}(x)\}^2\leqq 1$. Thus, by Theorem A we have $\{f^{(k)}(x)\}^2+\{f^{(k-1)}(x)\}^2\leqq 1$, or

$$\left\{\frac{1}{\lambda^k}\phi^{(k)}(x)\right\}^2+\left\{\frac{1}{\lambda^{k-1}}\phi^{(k-1)}(x)\right\}^2\leqq 1,\qquad k=1,2,3,\cdots.$$

Considering any fixed k and letting $\lambda\to N+0$, we obtain

$$\left\{\frac{1}{N^k}\phi^{(k)}(x)\right\}^2+\left\{\frac{1}{N^{k-1}}\phi^{(k-1)}(x)\right\}^2\leqq 1,\qquad k=1,2,3,\cdots.$$

To obtain conditions of equality Theorem C is directly applicable.

Finally we give another simple inference from Theorem A.

If $f_1(x)$ and $f_2(x)$ satisfy the conditions of Theorem A *for some integer n, then*

$$\left|\frac{d^k}{dx^k}f_1(x)f_2(x)\right|\leqq 2^{k-1},\qquad k=1,2,3,\cdots,n.$$

PROOF. By the Leibniz rule

$$\frac{d^p}{dx^p}(f_1'+if_1)(f_2'+if_2)=\sum_0^p C_{p,\nu}(f_1^{(\nu+1)}+if_1^{(\nu)})(f_2^{(p+1-\nu)}+if_2^{(p-\nu)})$$

where the constants $C_{p,\nu}$ are the binomial coefficients. Then by Theorem A

$$\sum_0^p C_{p,\nu}|f_1^{(\nu+1)}+if_1^{(\nu)}|\cdot|f_2^{(p+1-\nu)}+if_2^{(p-\nu)}|\leqq\sum_0^p C_{p,\nu}=2^p.$$

---

[6] M. Riesz, Jahresbericht der Deutschen Mathematiker-Vereinigung, vol. 23 (1914), pp. 354–368. C. J. de la Vallée Poussin, Comptes Rendus de l'Académie des Sciences, vol. 166 (1918), pp. 843–846. G. Pólya and G. Szegö, *Aufgaben und Lehrsätze aus der Analysis*, vol. 2, pp. 35, 90.

To separate real and imaginary parts we may write

$$(f_1' + if_1)(f_2' + if_2) = f_1'f_2' - f_1f_2 + i\frac{d}{dx}(f_1f_2).$$

Differentiating this last expression p times and taking absolute magnitudes, we obtain

$$\left\{\frac{d^{p+1}}{dx^{p+1}}(f_1f_2)\right\}^2 + \left\{\frac{d^p}{dx^p}(f_1'f_2' - f_1f_2)\right\}^2 \leqq (2^p)^2,$$

which is more than we set out to prove. The functions $f_1(x) = \sin(x+\alpha_1)$ and $f_2(x) = \sin(x+\alpha_2)$ show that our constant is the "best possible."

PURDUE UNIVERSITY AND
STANFORD UNIVERSITY

ON THE CARATHÉODORY CONDITION FOR UNILATERAL VARIATIONS[1]

JULIAN D. MANCILL

The two formulations and proofs of the Carathéodory condition in the calculus of variations given by Graves[2] do not necessarily apply to the case when the minimizing curve may have arcs in common with the boundary of the region of admissible variations.[3] The purpose of this note is to show how his first formulation and proof can be modified so as to be applicable to unilateral (one-sided) variations in the plane.

An admissible curve

$$E_0: \quad x^\alpha = x^\alpha(t), \qquad t_0 \leqq t \leqq T, \alpha = 1, 2, \cdots, n,$$

which minimizes the integral

$$J = \int_{t_0}^{T} F(x_1, \cdots, x_n, x_1', \cdots, x_n')dt \equiv \int_{t_0}^{T} F(x, x')dt$$

in the class of all admissible curves joining two fixed points x_0 and X in space of n dimensions ($n>1$), must satisfy certain well known con-

[1] Presented to the Society, December 29, 1939.

[2] *Discontinuous solutions in space problems of the calculus of variations*, American Journal of Mathematics, vol. 52 (1930), pp. 13–19.

[3] Cf. Mancill, *The minimum of a definite integral with respect to unilateral variations*, Contributions to the Calculus of Variations, 1933–37, University of Chicago, 1937, p. 121, condition C_{3C}.

ditions.[4] Suppose that E_0 is an extremaloid joining the points x_0 and X and having corners at $x_1, \cdots, x_m$, and suppose that the function F_1 is different from zero along E_0 including both sides of corners. If the function

$$\Omega_0(x, p^-, p^+) = (\partial F^+/\partial x^\alpha)p^{\alpha-} - (\partial F^-/\partial x^\alpha)p^{\alpha+},$$

where $p^{\alpha-}$ and $p^{\alpha+}$ represent the direction cosines of the tangents to E_0 at the corners, is different from zero at each corner of E_0, there exists an $(n-1)$-parameter family

$$x^\alpha = \phi^\alpha(t, a) \tag{1}$$

of extremaloids defined for $t_0-\delta \leqq t \leqq T+\delta$, passing through the point x_0 for $t=t_0$, and containing[5] E_0 for $a^i=a_0^i$, $t_0 \leqq t \leqq T$. We shall now prove the following form of the Carathéodory condition:

THEOREM. *Let E_0 be an extremaloid joining the points x_0 and X and minimizing the integral J, satisfying the following conditions:*

(1) *E_0 is positively strong;*[6]

(2) *$\Omega_0 \neq 0$ at the corners on E_0;*

(3) *E_0 is uniquely determined by each of its elements (x, x');*[7]

(4) *the determinant*

$$D(t, a_0) \equiv \left| \frac{\partial \phi^\alpha}{\partial a^i} - \frac{\partial \phi^\alpha}{\partial t} \right|_{a=a_0}$$

does not vanish at the corners on E_0.

Then the determinant $D(t, a_0)$ does not change sign at the corners on E_0.

Suppose that the determinant $D(t, a_0)$ changes sign at a corner x_k where $t=t_k$. Also, suppose that the corners on all extremaloids of the family occur for fixed values of the parameter[8] t. Consider the one-parameter family of curves E_u constructed by Graves.[9] The value of the integral J taken along E_u is a function $J(u)$ whose derivative near $u=t_k$ is

$$J'(u) = -\mathcal{E}(\phi^{\alpha-}, \phi'^{\alpha-}, x'^{\alpha+}) \equiv -\mathcal{E}[\bar{t}(u), \bar{a}(u), x'^{\alpha+}(u)]. \tag{2}$$

The second derivative of $J(u)$ at $u=t_k$ is

---

[4] For a full statement of the problem and the classical necessary conditions, see Graves, loc. cit., pp. 2–3.

[5] Graves, loc. cit., p. 6.

[6] Cf. Graves, loc. cit., p. 9.

[7] Ibid., p. 11.

[8] Ibid., p. 7.

[9] Ibid., p. 14.

$$(3) \qquad J''(t_k) = -\left[(\partial F^+/\partial x^\alpha)x'^{\alpha-} - (\partial F^-/\partial x^\alpha)x'^{\alpha+}\right]\bar{l}'(t_k),$$

on account of the extremal property of E_0, and since the derivative of the function $\mathcal{E}[t_k, \bar{a}(u), x'^{\alpha+}(u)]$ is zero at $u=t_k$. This last statement follows from assumptions (1) and (3) of the theorem and the fact that we have assumed that the corners on all the extremaloids of the family occur for fixed values of the parameter t, and therefore the function $\mathcal{E}[t_k, \bar{a}(u), x'^{\alpha+}(u)] \geqq 0$ near $u=t_k$ and equals zero at t_k. By making use of the homogeneity property of the function F and the expression for $\bar{l}'(t_k)$ given by Graves,[10] the derivative (3) may be reduced to

$$(4) \qquad J''(t_k) = -\left[(x'^{\alpha-}x'^{\alpha-})(x'^{\alpha+}x'^{\alpha+})\right]^{1/2} \cdot \Omega_0(x, p^-, p^+)D(t_k+0, a_0)/D(t_k-0, a_0) < 0,$$

where $p^{\alpha-}=x'^{\alpha-}/(x'^{\alpha-}x'^{\alpha-})^{1/2}$ and $p^{\alpha+}=x'^{\alpha+}/(x'^{\alpha+}x'^{\alpha+})^{1/2}$. Consequently, $J(t_k)$ *is a maximum of the function* $J(u)$ *for* u *in a neighborhood of* t_k, since $J'(t_k)=0$ and $J''(t_k)<0$. Thus, E_0 could not minimize the integral J.

Let us now consider the problem of unilateral variations in the plane. In this case it can be shown that a Carathéodory condition as stated here applies to every arc of the minimizing curve *which is an arc of an extremaloid*. The condition may be stated in terms of the one-parameter family of extremaloids through any fixed point of such an arc. For such arcs of the minimizing curve as E_{23} immediately following a non-extremal arc E_{12} of the boundary of the region of admissible curves whose direction at the point 2 where it meets E_{12} is not directed towards the exterior side of E_{12} but is as shown in Fig. 1, the Carathéodory condition may be stated in terms of the one-parameter family of extremaloids whose members, at least those members on the admissible side of the minimizing curve, are tangent to the arc E_{12} of the boundary.[11] In order to insure the existence of the families of extremaloids just described, it is necessary to assume that the continuity and homogeneity properties of the integrand F hold in an extended region containing in its interior those arcs of the minimizing curve which are extremaloids in common with the boundary. It should also be pointed out that in the construction of the family of extremaloids tangent to E_{12}, it is necessary to assume that the Weierstrass $\mathcal{E}$-function is greater than or equal to zero along

[10] Loc. cit., p. 15.

[11] Mancill, loc. cit., pp. 105–107. For a discussion of the case when the extremaloid is directed towards the exterior side of E_{12} see pages 138–141. In this case the minimizing curve need not satisfy the corner conditions at the point 2.

E_{12} preceding 2.[12] The family of extremaloids is then composed of the family of extremals tangent to E_{12} in a neighborhood of and preceding the point 2, and the family of extremaloids containing the extremaloid E_{23} which is the *continuation* family of the extremals tangent to E_{12} near the point 2.

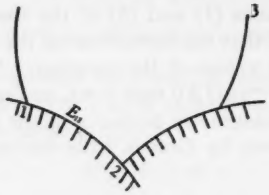

FIG. 1

In order to show how the proof of the theorem applies in the way just described, let (1) represent the one-parameter family of extremaloids through a fixed point of such an arc of the boundary. Let x_k be the *first corner beyond this fixed point.* Then the family of curves E_u used in the proof of the theorem exists for u near t_k and its members are admissible curves for u on one side of t_k. Therefore, the relation (2) holds here for u on the admissible side of t_k. The function $-\mathcal{E}[\bar{t}(u), \bar{a}(u), x'^{\alpha+}(u)]$ is defined locally and in a neighborhood of t_k. Its derivative at $u=t_k$ is $J''(t_k)$ as given in (4). Consequently, it follows from (2) that $J(t_k)$ is greater than $J(u)$ for all values of u on the admissible side of t_k and sufficiently near t_k. The same argument can now be made at each succeeding corner in turn.

UNIVERSITY OF ALABAMA

[12] Mancill, ibid., p. 107. For a proof of the Weierstrass condition along E_{12}, see Graves, American Mathematical Monthly, vol. 41 (1934), pp. 502–504.